Spatio-temporal analysis of wall-bounded turbulence: A multidisciplinary perspective via complex networks


Doctoral Dissertation
Doctoral Program in Aerospace Engineering (32^{nd} cycle)

Spatio-temporal analysis of wall-bounded turbulence

A multidisciplinary perspective via complex networks

Giovanni Iacobello

* * * * * *

Supervisor
Prof. Stefania Scarsoglio

Doctoral Examination Committee:
Prof. T. Karakasidis, Referee, University of Thessaly
Prof. A. Soldati, Referee, Technische Universität Wien
Prof. A. Braunstein, Politecnico di Torino
Prof. E. Ferrero, Università del Piemonte Orientale
Prof. M. Germano, Duke University

Politecnico di Torino
April 6th, 2020

I hereby declare that, the contents and organisation of this dissertation constitute my own original work and does not compromise in any way the rights of third parties, including those relating to the security of personal data.

. . . . . .

Giovanni Iacobello
Turin, April 6th, 2020

Summary

The aim of this Thesis is to propose a novel way to investigate wall-bounded turbulent flows, which relies on complex networks. The combination of fluid dynamics and complex networks theory results in an innovative and strongly multidisciplinary approach, which has been attracting growing interest. Differently from classical statistics, complex networks represent a suitable and powerful tool to capture and study the inter-connections between the elements of a discrete complex system. Therefore, the idea behind this work is to take advantage from the great developments of the recently established network science, to advance the level of information of classical statistics in the spatio-temporal characterization of wall turbulence.

Three main approaches are pursued to analyse turbulent flows in the view of (i) one-point time-series, (ii) Eulerian spatio-temporal fields and (iii) Lagrangian particle trajectories. About the first approach, one-point time-series are mapped in complex networks in which nodes correspond to temporal data, while links are established by means of the visibility algorithm. Here, for the first time the visibility algorithm is used to investigate a (numerically simulated) turbulent channel flow and a passive scalar plume in an experimental turbulent boundary layer. In both cases, particular efforts are paid to relate network metrics to the temporal structure of the signals. This fundamental issue has often been disregarded, since the meaning of the metrics has usually been interpreted only as a network feature rather than a signal feature. Results show that the visibility-based networks are able to quantify the presence of small fluctuations in the series, as well as the relative intensity of extreme events and their temporal occurrence. This information – usually hidden in classical statistics based on the signal PDF – is crucial, e.g., in the dispersion of contaminants in the atmospheric turbulent boundary layer, which is tackled in this Thesis via an experimental study.

The second approach concerns the Eulerian framework, and complex networks are exploited to explicitly emphasize high-correlation regions in a turbulent channel flow. In this case, nodes correspond to grid points and represent the volume of each cell in the computational domain. In this way, we extend the idea pursued in climate network analysis to weigh nodes on the corresponding surface areas. Links are activated if the two-point correlation of the velocity field is above a suitable

threshold, which is chosen in order to highlight large (positive and negative) correlation values. The main outcome is represented by the appearance of long-range connections – referred to as teleconnections – between distant near-wall locations. Teleconnections are found to be associated to an analogous response of the near-wall velocity field due to large-scale motions. The spatial texture of teleconnections is straightforwardly highlighted by the network links, which – differently to the classical average correlation – is able to retain the spatial information of high correlation coefficients.

Finally, in the third approach, fluid particle trajectories are exploited to study turbulent mixing in a channel flow by means of a time-varying weighted network. In this framework, nodes correspond to groups of particles and a spatial proximity criterion is employed to establish (weighted) links in the network, whose structure evolves in time. By doing so, we are able to clearly identify characteristic regimes of particle dispersion, as well as the interplay between advection and mixing on particle dynamics. The Lagrangian time-varying network approach, therefore, is able to capture the transient effects of the particle dynamics, differently from other works in which the full trajectories are exploited to activate links (thus loosing transient mechanisms).

For each of the three perspectives, complex networks reveal to be an effective tool to capture turbulence features that would be complicated to highlight through other techniques. More in details, complex networks are fully able: to inherit higher-order information of time-series (by highlighting the occurrence and relative intensity of extreme events); preserve the spatial information of correlation in all directions between different (even long-range) spatial locations; and to reveal local and global mixing effects of particle dispersion (thus providing in a unique framework a geometrical representation of particle dynamics over time). Based on present findings, the complex network approach can pave the way for a novel multidisciplinary research area that, beyond the proposed applications, can involve other turbulence configurations as multiphase and thermal or density inhomogeneous flows.

Acknowledgements

My first, great acknowledgement goes to my Supervisor, Prof. Stefania Scarsoglio, who let me the possibility to undertake the unique experience of a PhD. A special thanks also goes to Prof. Luca Ridolfi who, together with my Supervisor, closely followed my PhD journey. Seneca once stated that "*if one does not know to which port one is sailing, no wind is favourable*". By exploiting this effective metaphor, I would like to remark my gratitude to Professors Scarsoglio and Ridolfi for their invaluable support and wise advice, which always let me find *favourable winds* in the tumultuous seas of scientific research, thus fostering my growth as a young researcher.

I would like to extend my gratitude to Prof. Hans Kuerten, who hosted me in the Department of Mechanical Engineering at the Technical University of Eindhoven. His profound academic experience and invaluable numerical data availability have provided a turning point in my PhD.

I would also like to deeply thank Prof. Pietro Salizzoni. By hosting me in the Laboratoire de Mécanique des Fluides et Acoustique at the École Centrale de Lyon, I had the inestimable possibility to directly work on advanced experimental facilities, which has resulted in a very effective collaboration. I also express my gratitude to Massimo Marro for his support during the experiments in the wind tunnel.

I would like to profoundly thank Prof. Karakasidis and Prof. Soldati for serving as reviewers of my PhD Thesis, for their positive comments and appreciations, as well as their valuable suggestions that ultimately helped to improve my work.

Last, but not least, I express my deepest gratitude to everyone who has participated, directly or indirectly, to my PhD journey. I thank my family, that supported me throughout the PhD, and Anna who – more than anyone else – has always encouraged me during all the uncertainties I experienced as a young researcher. Finally, I thank Alberto, Alessia, Caterina, Francesco, Giuseppe, Marco and Maurizio for sharing the everyday adventures at the Politecnico di Torino, I say thanks to Vera for her kindness and useful suggestions during network conferences, as well as all the amazing friends that I met in Lyon, namely, Andrea, Chi Vuong, Cristina, Fabio, Fateh, Jheyson, Joy, Livia, Marcia, Matteo, Mehdi, Michel, Paul and Sofia.

Thank you all.

Contents

List of Tables

List of Figures

Nomenclature

NETWORK METRICS:

$\Gamma^n(i)$	n-th neighbourhood of node i (Pag. 30)
$\Gamma^n_c(i)$	n-th cumulative neighbourhood of node i (Pag. 30)
(a_x, a_y, a_z)	Ellipsoid semi-axis lengths along the three Cartesian directions (Pag. 53)
$\mathcal{A}_{i,j}$	Adjacency matrix (Pag. 30)
$\mathcal{C}(i)$	Clustering coefficient of node i (Pag. 34)
$\mathcal{D}_{i,j}$	Distance metric based on the correlation coefficient (Pag. 128)
$\mathcal{E}_p$	Total number of connections between particles (Pag. 153)
$\mathcal{F}_{g,q}$	Interaction weighting function between particles $\{g, q\}$ (Pag. 148)
$\mathcal{G}$	Mathematical representation of a graph (Pag. 30)
$\mathcal{K}(i)$	Degree of node i (Pag. 32)
$\mathcal{K}^{vw}(i)$	Volume weighted connectivity of node i (Pag. 33)
$\mathcal{K}^{vw}_{nn}(i)$	Average nearest-neighbours $\mathcal{K}^{vw}$ of node i (Pag. 36)
$\mathcal{L}$	Network average path length (Pag. 36)
$\mathcal{N}_e$	Total number of links in a network
$\mathcal{N}_v$	Total number of nodes in a network
$\mathcal{R}$	Subset of nodes in the network
$\mathcal{R}(\Gamma^n)$	Spatially-connected subset of nodes in the n-th neighbourhood (Pag. 117)

$\mathcal{R}(\Gamma_c^n)$	Spatially-connected subset of nodes in the n-th cumulative neighbourhood (Pag. 117)
$\mathcal{R}_L(\Gamma^1)$	Long-range region (Pag. 117)
$\mathcal{R}_S(\Gamma^1)$	Short-range region (Pag. 117)
$\mathcal{S}(i)$	Strength of node i (Pag. 33)
$\mathcal{T}$	Network transitivity (Pag. 34)
$\mathcal{W}_{i,j}$	Weight matrix (Pag. 30)
$\overline{d^w}_{x,y,z}(i)$	Weighted physical distance of node i along x, y or z directions (Pag. 37)
ϕ	Average peak occurrence (Pag. 37)
ϱ	Network edge density (Pag. 32)
$d^t(i)$	Mean link-length of node i (Pag. 37)
l_i	Particle level i (Pag. 146)
r	Network assortativity coefficient (Pag. 35)

GENERAL NOTATION:

$\bullet^*$	Time-series normalization with zero-mean and unitary standard deviation value
$\bullet^+$	Variable normalization in wall-units
$\delta_{i,j}$	Kronecker delta
$\langle\bullet\rangle$	Average over homogeneous directions and time
$\mathbb{I}$	Indicator (binary) function (Pag. 148)
$\mid\bullet\mid$	Absolute value (for scalars) or set cardinality
$\overline{\bullet}$	Arithmetic average
σ	Standard deviation function
$\mathrm{cov}\,[\bullet,\bullet]$	Covariance function
θ	Threshold value

Ku	Kurtosis function
$P(\bullet)$	Cumulative probability distribution
Sk	Skewness function

VARIABLES:

δ	Boundary layer thickness
ϵ_y	Wall-normal mixing coefficient (Pag. 144)
γ	Intermittency factor (Pag. 85)
κ	von Kármán constant
Λ_y, Λ_z	Average spread of passive scalar plume in directions y, z (Pag. 185)
$\lambda_{y,z}$	Instantaneous spread of passive scalar plume in directions y, z (Pag. 78)
$\mathcal{V}_\bullet$	Volume of a region in the domain
ν	Fluid kinematic viscosity
ω	Vorticity
ρ	Fluid density
τ	Total shear stress
$\widetilde{k}$	Normalized wavenumber
c	Passive scalar concentration
$C_{i,j}$	Correlation coefficient between signals in position i and j
D_s	Internal diameter of the passive scalar source (Pag. 182)
$d_{x,y,z}$	Physical (Euclidean) distance along x, y, or z
E	Power spectral density
F	Mean driving force per unit mass
f	Frequency
H	Channel half-height
h_r	Wall surface-roughness height (Pag. 182)

h_s	Height of the passive scalar source (Pag. 182)
h_s^*	True plume axis height (Pag. 185)
$L_{x,y,z}$	Total physical domain lengths along x, y and z
M_e	Mass flow rate of the passive scalar (Pag. 182)
M_t	Total mass flow rate of the passive scalar source (Pag. 182)
$N_{x,y,z}$	Number of physical domain grid points along x, y and z
p, p_t	Static and total pressure
Re	Reynolds number
$s(t_i)$	Generic time-series evaluated at time t_i
T	Total simulation/experiment time
t, t_i	Time variable, i-th time instant
u', v', w'	Streamwise, vertical, spanwise fluctuating velocity components
U, V, W	Streamwise, vertical, spanwise mean velocity components
u, v, w	Streamwise, vertical, spanwise velocity components
U_∞	Free-stream mean-velocity in the boundary-layer flow
u_τ	Friction velocity
x, y, z	Streamwise, vertical, spanwise directions
X_i, Y_i, Z_i	Coordinates in the x, y, z directions of point i

Chapter 1

Introduction

"When things get too complicated, it sometimes makes sense to stop and wonder: have I asked the right question?"

Enrico Bombieri

Fluid dynamics has always played a crucial role in aerospace engineering research. The rapid technological developments of the aeronautical industry pursued at the beginning of the last century, have largely benefited from the advances in the understanding of fluid-structure interaction. In practical applications, in fact, fluid flows interact with solid surfaces (e.g., an aircraft wing or a rotor blade), thus inevitably generating a thin boundary layer in which the effects of fluid viscosity are substantial. In the more general area of fluid dynamics, turbulent flows deserve peculiar attention, since they are associated with energy and power losses in transporting fluid (e.g., in pipes and channels) of moving object through the fluid (i.e., basically any transportation system, from aircraft to cars and ships). The importance of understanding turbulent flows – i.e., their origin, their spatio-temporal organization and their control – is corroborated by their appearance in almost all practical applications. Consequently, turbulent flows have always raised the interest of different scientific communities, either engineering, physical or mathematical.

Despite important progresses have been achieved in the last century about the characterization of turbulence, fundamental issues – e.g., concerning statistics scaling, coherent motion or energy consumption – are still the object of a worldwide active research. In last decades, in particular, thanks to the outstanding growth of computational technologies, a huge amount of data has been generated for validating existing theories and finding out unknown turbulence features. The large availability of data has allowed to tackle the intrinsic complexity of turbulence via a novel perspective, namely the data-driven approach has been established instead of the (classical) theoretical approach. With the aim to overcome the boundaries

of turbulence complexity, therefore, novel tools of analysis have been continuously developed to exploit such *big-data*.

Among other techniques, network science – i.e., the study of complex systems by means of complex networks – strikingly emerged in the last twenty years, providing a framework to investigate several real-world complex systems. Examples are social interactions, biological processes, trades and economic systems as well as transportation and computer science. The peculiarity behind the great success of network science is its ability to highlight non-trivial features of the whole complex system, through a texture of simple pairwise connections between the constitutive elements of the system. Although they were historically developed for the study of social behaviours, complex networks have recently found a breeding ground in technical applications (e.g., computer science, brain diseases and climate studies), providing a strongly multi- and inter-disciplinary lens on (among others) physical and engineering problems. As a consequence, a change of paradigm has come into being, from the classical social network analysis to a physics-based network approach.

The present Thesis aims to provide a basis for a novel perspective on the spatio-temporal characterization of wall turbulence, which fully takes advantage from the great potential offered by network science. In fact, although turbulence and network science have been singularly investigated, the combination of turbulent flows and complex networks results in a strongly innovative and multidisciplinary approach. Accordingly, complex networks are used in this Thesis to capture in a systematic way non-trivial features of turbulent spatio-temporal fields, such as the occurrence and relative intensity of extreme events in one-point signals, the emergence of strongly correlated regions displaced at distant spatial locations, as well as the interplay between mixing and advection in the dispersion of a set of fluid particles. These outcomes, which would be complicated to highlight by means of other techniques, are straightforwardly characterized by complex network tools. In light of this, the present work aims to cover different viewpoints on turbulence, namely time-series analysis as well as the Eulerian and Lagrangian perspectives. Complex networks can then be generated according to turbulence data availability, each representing a different scenario of analysis. Therefore, based on the findings reported in this work, the network-based approach can be effectively employed – beyond the proposed applications of wall-bounded turbulent flows – to other turbulence configurations, such as flows with thermal or density inhomogeneity, compressible flows, as well as multiphase flows.

The Thesis is organized as follows.

Chapter 2 comprises an introduction to turbulent flows with emphasis on wall-bounded turbulence. The main features of wall turbulence are outlined and its wall-normal structure is described in terms of the characteristic physical parameters. A

brief survey of the main open issues and achievements in different sub-areas of wall turbulence is also reported, thus providing a literature background for the network-based applications pursued in this work.

Complex networks are introduced in Chapter 3. The principal applications as well as the historical background that has led to the development of a science of networks are briefly presented. An overview of the main network definitions is also reported in Chapter 3, which delineates the features and the mathematical formulation of the network metrics exploited in this work.

The three approaches to map turbulent spatio-temporal fields into complex networks – i.e., time-series, Eulerian and Lagrangian viewpoints – are illustrated in Chapter 4. An up-to-date review of the network-based applications to vortical and turbulent flows is reported, by highlighting the difference and analogies between the methodologies that have been developed so far and the approaches proposed in the present work.

Results for each of the three approaches are shown and described in Chapter 5, Chapter 6 and Chapter 7. Specifically, the time-series analysis of turbulence data is reported in Chapter 5, in which the visibility graph method is exploited to map time-series into complex networks. Two wall-bounded turbulence setups are investigated: (i) a numerically simulated turbulent channel flow from the Johns Hopkins turbulent database (JHTDB) that is publicly available online; (ii) experimental data of a passive scalar plume in a turbulent boundary layer. Experimental data are measured in the Laboratoire de Mécanique des Fluides et d'Acoustique in collaboration with Professor P. Salizzoni (École Centrale de Lyon).

The Eulerian approach to turbulence is delineated in Chapter 6, in which a numerically simulated turbulent channel flow is exploited to build *spatial networks* by thresholding velocity-based correlation coefficients. Data from the numerical simulations are obtained in collaboration with Professor J.G.M. Kuerten (Eindhoven University of Technology), for different temporal windows and two Reynolds numbers. The numerical code and runs are acknowledged to Prof. Kuerten.

Chapter 7, instead, shows the results of turbulent mixing analysis in a Lagrangian viewpoint. Particle trajectories extracted from a numerically simulated turbulent channel flow are exploited to build a complex network that evolves in time, to capture transient effects of mixing and advection. Similarly to the Eulerian approach, even in this case the numerical simulation exploited to extract particle trajectories is performed in collaboration with Professor J.G.M. Kuerten (Eindhoven University of Technology). The numerical code and runs are acknowledged to Prof. Kuerten.

Finally, Chapter 8 summarizes the results obtained in this work and draws some concluding remarks.

Chapter 2

Turbulent flows

"In any case there is considerable 'turbulence' in the attempts to define what is turbulence"

Arkady Tsinober

Fundamental features of turbulence are introduced in this Chapter, by providing an overview of the main achievements obtained so far, current open issues as well as typical tools for turbulence analysis. In particular, wall-bounded turbulent flows are focused as they represent the object of this work. From an engineering point of view, in fact, wall-bounded turbulence plays a crucial role for technological, energetics and environmental developments.

The Chapter is organized as follows. In Section 2.1 the main general features of turbulence are reported. Section 2.2 comprises the description of wall-turbulence features as a function of the distance from the wall. The main theoretical and practical achievements and challenges in wall-turbulence are reported in Section 2.3; specifically, emphasis is given to scaling and coherent motion analyses (Section 2.3.1) as well as the dispersion of passive scalar in wall-turbulence (Section 2.3.2).

2.1 Introduction

The possibility to understand and predict the behaviour of fluids has always been a crucial aspect in scientific research, due to the importance of gases and liquids in everyday life. By looking at the fluid motion in nature, it clearly appears that – except few rare cases – the behaviour of gases and liquids is characterized by a chaotic, complicated and unpredictable motion, which is referred as *turbulence*. Despite the ubiquitous nature of turbulent flows, their definition is still lacking due to the huge complexity intrinsic to turbulence. Therefore, in order to understand

turbulence physical features, the typical approach is to show examples of turbulent flows, as reported in the firsts chapters of several well-established textbooks concerning turbulence (among others, see [1–5]). The atmospheric and oceanic flows, the wake past vehicles and aircraft, the surface of stars (e.g., the Sun), as well as the flow in pipes, clouds, rivers and chemical reactors are just few examples of the plethora of applications in which turbulence is involved. In the case of atmospheric flows (e.g., see Figure 2.1), turbulence understanding is fundamental, for instance, for climate analyses, dispersion of contaminants as well as cloud formation [6]. Besides, turbulence plays a major role in engineering applications, such as aeronautics, automotive, civil and mechanical engineering, where the understanding of the effects induced by the interaction turbulence-structure (e.g., drag or lift forces) is fundamental.

Despite the lacking of a conclusive definition, some characteristic features of turbulent flows have been outlined so far by relying on everyday experience [1]. The irregularity (i.e., the chaotic motion) and the intense vorticity (i.e., the swirling behaviour) are undoubtedly the most apparent features of turbulence. In this respect, non-linearity in the equations of motion (i.e., Navier-Stokes equations) plays a key role in generating turbulence as a chaotic and (apparently) random process. One of the most important consequences of this chaotic motion is the ability of turbulence to enhance mixing – that is the transport of momentum, heat and mass – which is much higher than molecular diffusion [7, 8]. Moreover, turbulence produces dissipation, as the turbulence kinetic energy is converted into internal energy of the fluid by means of viscous shear stress. It directly follows that without a supply of energy, turbulent processes rapidly decay. Specifically, the distribution of kinetic energy is not uniform in a turbulent flow, but it depends on the considered *scale*. Indeed, turbulence possesses a wide range of coexisting and interacting scales in space and time: they range from the largest scales, which are related to the geometry in which the flow develops, to the smallest ones, which are the scales where dissipation acts [4]. The typical visualization of a highly turbulent flow, in fact, is characterized by the presence of several vortical structures of different nominal sizes (e.g., see Figure 2.1). The degree of separation between large and small scales is regulated by the Reynolds number: for increasing values of the Reynolds number, the size of the smallest scales decreases (as the largest one are usually fixed by boundary conditions). Since turbulent flows in nature display high values of the Reynolds number, the smallest scales are many order of magnitude lower than the largest scales: this turns out to be a crucial issue for turbulence simulations and experimental reproducibility. Finally, turbulence involves three-dimensional (3D) processes, since vortex-stretching (that is a fundamental turbulence maintaining mechanism) requires 3D vorticity fluctuations [1]. However, it should be noted that in many physical systems – such as large scale circulation in atmospheric and oceanic flows – the two-dimensional approximation takes relevant information on the dynamics of the fluid [9].

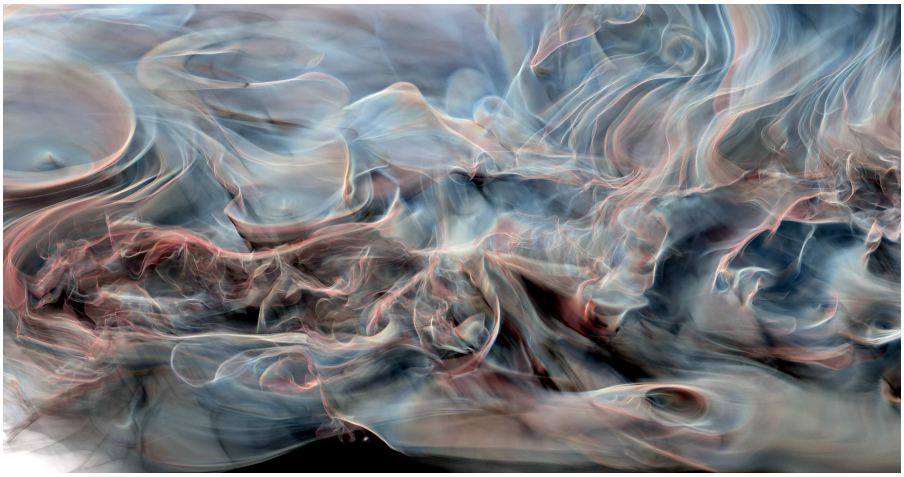


Figure 2.1: A large-scale visualization of a numerically-simulated turbulent atmospheric air-flow [10].

Since there are many ways in which a turbulent flow can arise with its specific features (e.g., homogeneity, isotropy, body forces, compressibility, magnetic fields and so on) there is not a unique problem about turbulence to address, but several inter-related problems that depend on the specific setup [5, 11]. In this work, wall-turbulence is focused because it plays a crucial role in industrial and engineering applications, as any product (e.g., pipelines, aircraft, ships, vehicles, combustors) have to interact with turbulence through a solid surface. Moreover, the atmospheric air motion is also a bounded turbulent flow as it must interact with the ground, including the effects of land morphology and urbanization. As a paradigmatic example, it is estimated that about a quarter of the energy spent by industry and transportation is used to move fluids in pipelines or to overcome the drag force in vehicles [12]. Furthermore, other estimates indicate that a 5% reduction of the transportation energy loss in the US would have an environmental impact comparable to a doubling of wind energy production [13]. The investigation of wall turbulence, therefore, has a direct impact in improving both environment-related technologies and industrial performances.

2.2 Structure and features of wall-turbulence

In this section, the wall-normal structure of wall-bounded turbulent flows is described by highlighting its main characteristics. Although many features here reported are also shared with other geometries (e.g., pipe flows), we focus only on two setups: turbulent channel flow and turbulent boundary layer. In both configurations, we investigate fully-developed turbulence, namely we focus on spatial regions that are far enough from the flow inlet. In what follows, we indicate the

streamwise, wall-normal and spanwise directions as x, y and z, respectively. The corresponding velocity components in each direction are indicated as (U, V, W) and (u', v', w') for the mean velocity and their fluctuations, respectively. In particular, the mean velocity is evaluated by averaging in time and over homogeneous directions. For both configurations analysed, $V = W = 0$ as the mean flow direction is streamwise. Furthermore, we focus only on flows that are statistically stationary and homogeneous in the spanwise and streamwise directions, with no-slip condition at the wall.

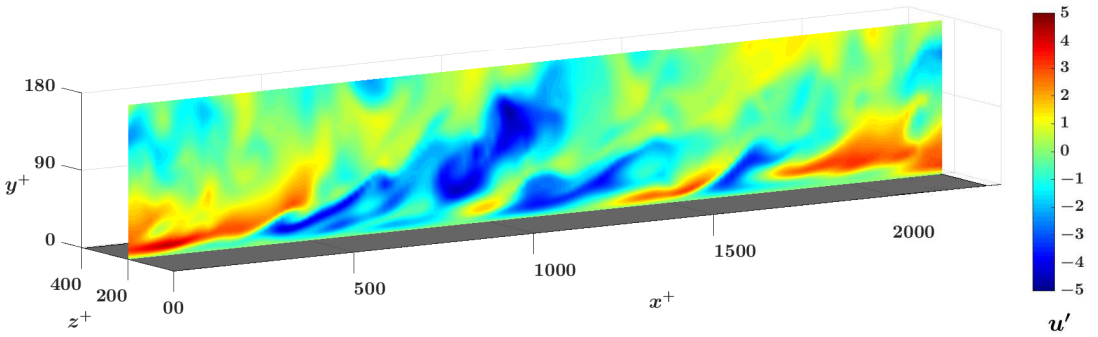


Figure 2.2: 3D view of a channel flow at $Re_\tau = 180$ (see Appendix B.3). A view of a two-dimensional section at $z^+ = 200$ is also illustrated where colors refer to the streamwise velocity fluctuation, u'. The mean flow is from left to right, i.e. for increasing x^+ values.

To characterize the turbulent channel flow and the turbulent boundary layer, two Reynolds numbers are defined, respectively, as

$$Re_\tau \equiv \frac{H u_\tau}{\nu}, \qquad Re_\delta \equiv \frac{\delta U_\infty}{\nu}, \tag{2.1}$$

where u_τ is the friction velocity, ν is the kinematic viscosity of the fluid, H is the half-channel height, while δ and U_∞ are the thickness and free-stream velocity in the boundary-layer flow, respectively. In particular, u_τ takes into account the friction at the wall and is defined as $u_\tau \equiv (\tau_w/\rho)^{0.5}$, where ρ is the fluid density and τ_w is the wall shear stress. Accordingly, wall-units are defined in terms of u_τ and ν with the superscript $+$: lengths and times are normalized by dividing ν/u_τ and ν/u_τ^2, respectively, while velocities are simply normalized by u_τ. As an example of setup, Figure 2.2 shows a three-dimensional view of a numerically simulated turbulent channel flow at $Re_\tau = 180$, where colors in a (x^+, y^+)-plane are the values of the streamwise velocity fluctuation, u', in that section.

Since the presence of the wall introduces anisotropy in the flow along the y direction, the vertical coordinate in wall-units, $y^+ = y u_\tau/\nu$, is used to discriminate different regions in the flow. Accordingly, the parameters that are involved in the

wall-normal layer characterization of the flow can be classified based on the distance from the wall. Near-wall parameters (i.e. very close to the wall) are represented by the viscosity and the friction velocity; on the other hand, outer-flow parameters are related to the geometry (via H or δ) and the external velocity (that is the centreline velocity, U_c, for the channel flow and the free-stream velocity, U_∞, for the boundary layer). In the following, in order to simplify the description, we only focus on the channel flow configuration; however, similar findings are valid for the turbulent boundary layer if δ and U_∞ are used in place of H and U_c, respectively.

The relative effects of near-wall and outer-flow parameters on the flow features discriminate different regions along the wall-normal direction. Such regions are basically related to the behaviour of the mean velocity, U, and the shear stress, τ [14]. In particular, τ is made up of two contributions

$$\tau(y) = \rho\nu\frac{\mathrm{d}U}{\mathrm{d}y} - \rho\langle u'v'\rangle, \quad \text{or} \quad \frac{\tau(y^+)}{\rho u_\tau^2} = \frac{\mathrm{d}U^+}{\mathrm{d}y^+} - \frac{\langle u'v'\rangle}{u_\tau^2}, \tag{2.2}$$

where $-\rho\langle u'v'\rangle$ is the Reynolds shear stress and $\rho\nu\mathrm{d}U/\mathrm{d}y$ is the viscous stress. While the mean velocity gradient (in wall-units) takes into account the contribution to the shear stress due to viscosity, the Reynolds shear stress quantifies the contribution due to turbulent velocity fluctuations. In Eq. (2.2), $\langle\bullet\rangle$ indicates averaging over homogeneous directions and time, while the minus sign is due to the fact that $\langle u'v'\rangle < 0$. It should be noted that, for the channel flow $\tau = \tau(y^+)$, while τ depends also on the streamwise location, x, in a boundary layer. However, at fixed x-coordinate, a similar reasoning holds for the channel and boundary layer flows.

In order to illustrate the characteristic regions in a wall-bounded turbulent flow, in Figure 2.3 we report the behaviour of the mean velocity (Figure 2.3(a)) and the two contributions to the total shear stress (Figure 2.3(b)), for a numerically simulated turbulent channel flow at $Re_\tau = 950$. From the point of view of the mean velocity, the *inner layer* is defined as the region for $y/H \ll 1$ (at high Reynolds numbers) in which U^+ depends only on near-wall (viscous) parameters. Specifically, the inner layer is conventionally limited to $y/H = y^+/Re_\tau < 0.1$ [4]. On the contrary, the *outer layer* is defined as the region in which the effect of viscosity on U^+ is negligible, and it ranges for $y^+ > 50$ [4]. The region for $y^+ < 50$, therefore, is indicated as *viscous wall region*, as the viscous contribution to the total shear stress is significant (at $y^+ \approx 50$ the viscous shear stress is about 10% of the total, see Figure 2.3(b)). More in detail, four regions are identified by moving upwards from the wall (see Figure 2.3(a)):

- very close to the wall, for $y^+ \lesssim 5$, the viscous stress is much more significant than the Reynolds shear stress (see Figure 2.3(b)), so that $\mathrm{d}U^+/\mathrm{d}y^+ \approx 1$ and, as a result, $U^+ = y^+$. The region for $y^+ \lesssim 5$ is then referred as *viscous sub-layer*;

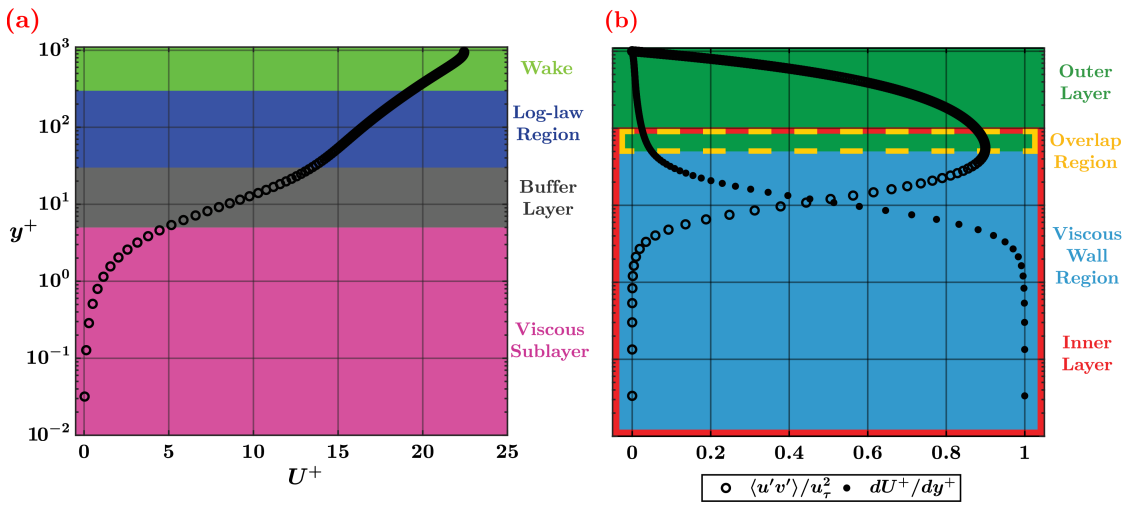


Figure 2.3: (a) Mean velocity profile, $U^+(y^+)$, for a numerically simulated channel flow at $Re_\tau = 950$ (see Appendix B.5). (b) Wall-normal behaviour of the Reynolds shear stress, $\langle u'v'\rangle/u_\tau^2$, and the mean velocity gradient, $\mathrm{d}U^+/\mathrm{d}y^+$. The limits of the highlighted regions are reported in the main text.

- at higher distances from the wall, the viscous stress decreases while the (turbulent) Reynolds shear stress increases (see Figure 2.3(b)): the dependence of $U^+(y^+)$ on viscosity is then negligible and the mean velocity follows a *log-law* profile:

$$U^+ = \frac{1}{\kappa}\ln y^+ + B, \tag{2.3}$$

  where $\kappa \approx 0.41$ is the von Kármán constant, and B is a constant that depends on wall roughness (for smooth walls, $B \approx 5.2$) [1, 4]. The log-law holds in the *log-law region*, which conventionally ranges from $y^+ \gtrsim 30$ up to $y/H = 0.3$;

- in between the viscous sub-layer and the log-law region, the mean velocity profile joins up the relations $U^+ = y^+$ and $U^+ = \ln y^+/\kappa + B$ (see Figure 2.3(a)). This intermediate region in the range $5 < y^+ < 30$ is called *buffer layer* and the Reynolds shear stress and viscous stress are of comparable intensity in this layer (see Figure 2.3(b)). It is important to underline that the maximum of the production of turbulence kinetic energy, $-\langle u'v'\rangle \mathrm{d}U/\mathrm{d}y$, is found in the buffer layer; specifically, the y^+ location of maximum production corresponds to the y^+ value in which $-\langle u'v'\rangle = \mathrm{d}U/\mathrm{d}y$, that is at about $y^+ \approx 12$ [4];

- for $y^+ > 0.3$, the flow behaviour depends on the outer-flow parameters as well as the geometry, and such region is referred as *wake* or *core* region [14].

In the inner layer (i.e., for $y/H \ll 1$), the behaviour of the mean velocity can be reduced to an expression $U^+ = f_w(y^+)$ that is universal (it does not depend on the flow setup) and that is referred as the *law of the wall* [4]. Specifically, in the viscous sub-layer, $f_w(y^+) = y^+$ while in the log-law region $f_w(y^+)$ is given by Eq. (2.2). On the contrary, in the outer layer, the mean velocity follows the so called *velocity defect law*, that is $(U_c^+ - U^+) = f_H(y/H)$, because U^+ depends on y/H as the effect of viscosity is negligible. As shown in Figure 2.3(a) for a channel flow, for $y/H > 0.3$ (i.e., $y^+/Re_\tau > 0.3$), the velocity defect law follows quite closely a logarithmic behaviour as in the log-law region. However, it should be emphasized that the arguments used for obtaining the Eq. (2.2) are not valid for $y/H > 0.3$. Additionally, while the function f_w is universal, the function f_H depends on the flow. For instance, the deviation from a logarithmic law of f_H are much more substantial for a turbulent boundary layer than in a channel flow [4]. Finally, since the limit of inner layer in terms of y^+ depends on the Reynolds number, it follows that (for high Reynolds numbers) there is an *overlap region* (also called *inertial sublayer*), $50 < y^+ < 0.1Re_\tau$, between the inner and the outer layer [1]. The presence of the overlap region (see Figure 2.3(b)) implies that the considerations for both the inner and outer layer are valid. In the overlap region, therefore, the mean velocity profile is commonly supposed to follow a logarithmic law of the form

$$U^+ = \frac{1}{\kappa} \ln\left(\frac{y}{H}\right) + B', \tag{2.4}$$

where B' is a constant depending on the flow and (if present) on wall-roughness [4, 15]. However, the behaviour of the mean velocity in the overlap region is still debated, as it will be discussed in Section 2.3.1.

Although the wall-normal structure of wall turbulence is basically associated with the behaviour of the mean velocity and shear stress, the statistics for velocity fluctuations are also reported. Indeed, velocity fluctuations account for the kinetic energy of turbulence, so that their behaviour is essential for the wall-normal characterization of turbulence. Figure 2.4 shows the profiles of the Reynolds stresses, $\langle u'^2 \rangle$, $\langle v'^2 \rangle$ and $\langle w'^2 \rangle$, as a function of y^+. $\langle u'^2 \rangle$ and $\langle w'^2 \rangle$ increase quite rapidly from the wall (where all velocity are zero) reaching a peak within the viscous wall region, $y^+ < 50$ (see Figure 2.4(a)). On the other hand, $\langle v'^2 \rangle$ increases rather slowly from the wall, which "constrains" the vertical motion of the fluid. It should be emphasized that the value of the peak of $\langle u'^2 \rangle$ increases with the Reynolds number for channel and boundary layer flows, as it increases the influence of (outer) large scales on the near wall region [16]. For high values of y^+, instead, the Reynolds stresses approach the same value as the anisotropy of the flow is less severe away from the wall. However, the intensity of the streamwise fluctuations, $\langle u'^2 \rangle$, is much higher than the other two components, especially close to the wall. Furthermore, as shown

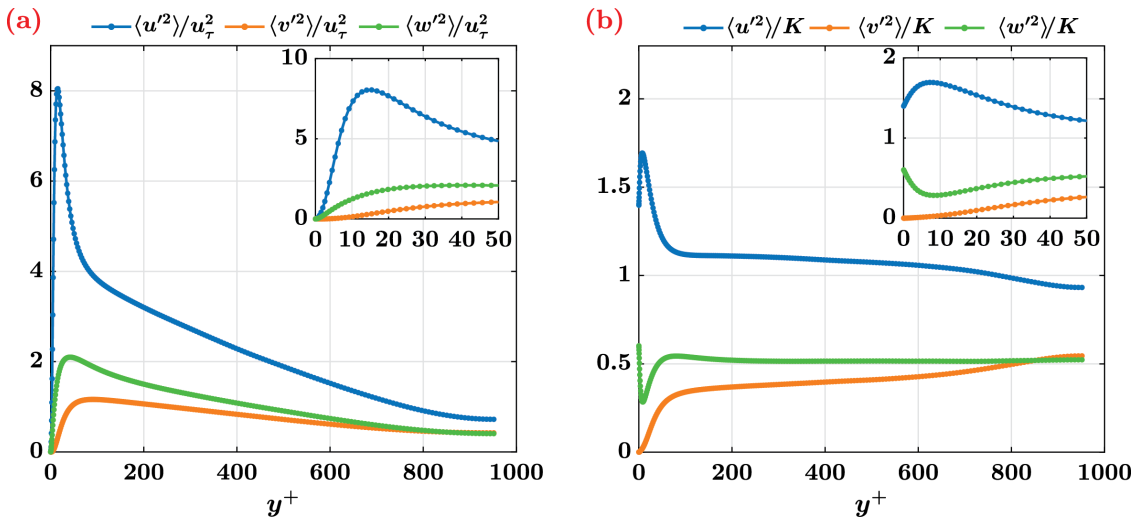


Figure 2.4: Profiles of the Reynolds stresses as a function of the wall-normal coordinate, y^+, normalized (a) in wall-units and (b) by the turbulent kinetic energy, K. The insets show the zoomed profiles for $y^+ \leq 50$. Data are from a numerically simulated channel flow at $Re_\tau = 950$.

in Figure 2.4(b), the streamwise velocity component accounts for most of the turbulent kinetic energy at all y^+. This makes the (mean and fluctuating) streamwise velocity component a fundamental variable in the wall turbulence characterization.

2.3 Achievements and challenges: an overview

The intrinsic complexity of turbulence has challenged the researchers to find out footprints of universality as well as the presence of coherence in such a complicated motion. In order to disentangle turbulence complexities, hence, the statistical approach has necessarily arisen, mainly for two reasons. First, the combination of non-linearity and acute sensitivity to perturbations in the initial and boundary conditions, make turbulent flows unpredictable in details. This has induced the scientific community to look at the turbulence from a probabilistic view. In fact, although the details of any turbulence realization are unique in practice, the statistical properties of turbulent flows under (almost) the same conditions are observed to persist [2]. Second, even when detailed information is available, any observed phenomenon concerning turbulence has to occur often enough, namely it has to be relevant for the turbulence dynamics from a statistical viewpoint [12].

Historically, the Reynolds decomposition of physical quantities in a time-averaged and a fluctuating component has provided a first notable achievement for the statistical analysis of turbulence [17]. Besides, the seminal papers of Taylor on particle

dispersion [18, 19] as well as the works of Kolmogorov on the characterization of turbulence at different scales [20, 21], have further contributed to develop a statistical theory of turbulence. These works (first by Reynolds, and later by Taylor and Kolmogorov) have provided in the last century a grounded framework for fundamental developments on turbulence in a very wide range of applications. For example, the Reynolds approach to Navier-Stokes equations (see Eq. (B.1) and (B.2)) – namely the Reynolds-average Navier-Stokes simulation (RANS) – has revealed to be a powerful tool for practical investigations. Nowadays, it is widely employed in computational fluid dynamics of engineering applications, as the RANS approach is able to capture essential (average) features of turbulence with reasonable computational costs [22].

The statistical approach to turbulence has gone side by side to the technological developments on flow simulation and experimental measurements. The former-developed single-point measurement techniques – e.g., Hot-Wire Anemometry (HWA) or Laser Doppler Velocimetry (LDV) – have strongly required a statistical view of turbulence. In fact, while numerical simulations and multi-dimensional experimental tools provide simultaneous information on the flow, single-point measurements performed at different locations are reasonably meaningful only from a statistical view since the simultaneousness is lost [12]. On the other hand, structural information about turbulence (i.e., two- and three-dimensional flow features such as visualizations of coherent structures) has been established only after the developments of direct numerical simulations (DNS) and experimental tools as particle image velocimetry (PIV) [12, 23]. For example, structural information about the region very close to the wall in a turbulent boundary layer and in a pipe flow was only given experimentally in 1960s by *Kline et al.* [24] and *Corino and Brodkey* [25], respectively. In particular, the visualizations of near-wall streaks in a turbulent boundary layer (see Ref. [24]) has considerably advanced the possibility to understand the mechanisms of wall-turbulence. For numerical simulations, accurate DNSs were mainly developed in the 1970s for homogeneous turbulence [26] and in the late 1980s for inhomogeneous flows [27], thanks to the increased computing availability. Nowadays, by exploiting the exponential boost of computational power, great efforts have been paid to the development of state-of-the-art techniques, such as machine learning and network science. For instance, machine learning has started to make an impact on turbulence modelling and analysis, thus providing an example of multidisciplinary approach to turbulence research (e.g., see Ref.s [28–30]).

From the picture outlined so far, it clearly emerges that novel theoretical and practical approaches are continuously required to shed light on the complex dynamics of turbulence. This topic was raised in the enlightening discussion "*Challenges in Turbulence in the Twenty-First Century–What Problems Should We Focus on in the Next 20 Years?*" [31], in which several authors indicate that turbulence research would get benefits by inheriting insights from other disciplines. Therefore,

by relying on state-of-the-art technologies, both novel experimental and numerical techniques are necessary to advance the level of information on turbulent flows. In this work, we focuses on the analysis of wall-bounded turbulence via complex networks, with emphasis on fully-developed turbulent channel flows and zero-pressure gradient turbulent boundary layers. In particular, the two typical points of view – namely the statistics scaling and the coherent structure characterization – are considered. The importance of scaling is related to the possibility to extract universal features of turbulent flows, with practical implications on flow predicative capability; on the other hand, the structural view (which is inter-related to scaling) is based on the possibility to describe the complex turbulent dynamics in terms of organized motion [14, 32]. We also highlight the effect of turbulence on the dispersion of a passive scalar, with emphasis on the role of the wall-induced inhomogeneity.

2.3.1 Wall-bounded flows: coherent structures and scaling

Statistical analyses of wall turbulence have mainly been focused on two interrelated topics: scaling and structures. While the former refers to the scaling exponents of typically investigated statistics (e.g., profiles of mean and fluctuating velocity, spectra, correlation and structure functions), the latter indicates the characterization of organized motions [33]. For both themes, a crucial role is played by the Reynolds number. In fact, as mentioned in Section 2.1, the higher the Reynold number the wider the range of turbulent scales. This affects the scaling properties (as statistics might change or not directly scale with the Reynolds number) and the features of coherent structures (as at higher Reynolds numbers larger structures appear, interacting with smaller ones). It should be noted that the presence of wall-roughness can affect turbulence features even far from the wall. In fact, the presence of small roughness elements barely influences the mean velocity profile except in the near wall region; however, roughness considerably affects the Reynolds shear stress, as well as the geometrical features of large scale coherent structures [34].

Statistics scaling

From the scaling perspective, one of the most debated issues in wall-turbulence concerns the asymptotic behaviour at high Reynolds numbers of the mean velocity profile, as well as high-order statistics (e.g., Reynolds stresses) [11, 14, 35]. The mean velocity profile (in wall-units) depends on two coefficients (see Eq. (2.2)): one related to wall roughness and setup geometry (e.g., channels or boundary layers), while the other is the von Kármán constant. The values of both these coefficients (in particular of the von Kármán constant) are still debated as only recently information from high Reynolds number turbulence are acquired to provide accurate estimations [33]. The importance of accurately evaluating such coefficients is related to the appearance of their values in turbulence models (such as the Spalart-Allmaras

or for RANS), thus affecting practical applications, e.g. for drag reduction [16]. This issue is very important for turbulent boundary layers, as the wall friction is more difficult to measure with high accuracy than for channel flows (where friction is estimated from the streamwise pressure gradient) [4].
Furthermore, it is traditionally assumed that the mean velocity in the overlap region follows a logarithmic profile as a function of the distance from the wall. Based on experiments and numerical simulations at high Reynolds numbers, other scaling theories have been proposed, as a power-law scaling of the mean velocity instead of a log-law scaling, with serious implications for turbulence modelling [16, 33, 36]. Very recently, however, it has been reported that physical constraints impose a universal log-law structure [35, 37], and universal agreement for different geometries emerges as well [38]. To get more insights into the wall-normal scaling, the streamwise velocity fluctuation, $\langle u'^2 \rangle$, has also been considered. According to the Townsend-Perry model of self-similar wall-attached eddies, $\langle u'^2 \rangle$ is expected to follow a logarithmic profile in the same region in which U is expected to be logarithmic [16]. Therefore, the behaviours of mean and fluctuating streamwise velocity are used as a diagnostic for log-law scaling in wall turbulence.

Another recent issue about scaling at high Reynolds numbers concerns the appearance of a peak in pre-multiplied spectrum of the streamwise velocity at low wavenumbers. This peak has been associated with the presence of very large scale motions (VLSMs) – also referred as superstructures – in the flow [12, 16, 33, 36]. VLSMs are organized structures with characteristic size of the order of ten times the characteristic scale in the flow (e.g., the boundary layer thickness or the channel height) and carrying large fractions of turbulence energy [36].

Coherent structures

Very large scale motions represent a remarkable example of coupling between scaling theories and coherent structures, as the behaviour of the statistics (e.g., pre-multiplied spectra) are directly related to the presence of organized (structural) motion in the flow. Coherent structures are very important as they provide a simplification of the description of the turbulence dynamics and are fundamental for turbulence control strategies [32]. Accordingly, several techniques have been employed to extract such coherent behaviour. It is worth to note that coherent structures have to possess both spatial and temporal coherence, namely they have to persist for enough long time [39].

In wall turbulence, the characterization of coherent structures depends on the wall-normal distance. In the near wall region, the important structures are *streaks* and *quasi-streamwise vortices* (or *rolls*) [12, 32]. Streaks are streamwise-elongated structures of negative streamwise velocity fluctuations (see Figure 2.5), with an average spanwise separation of $z^+ \sim 100$ and size $x^+ \sim 10^3 - 10^4$ and $z^+ \approx 50$ in the spanwise and streamwise direction, respectively [12]. Low-speed streaks

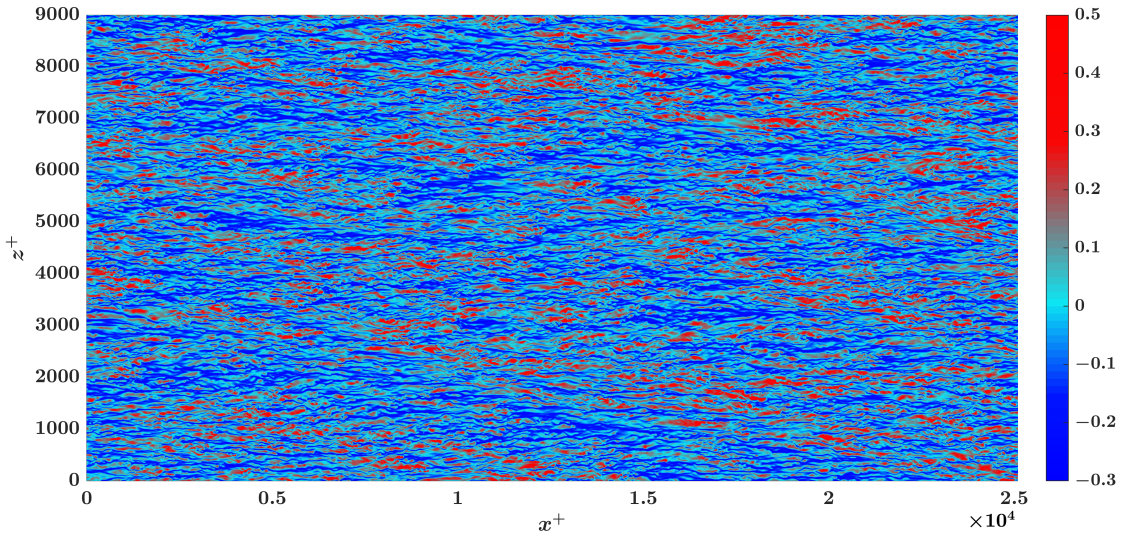


Figure 2.5: Streamwise velocity fluctuation in a wall-parallel plane at $y^+ \approx 10$. DNS data are extracted from the Johns Hopkins Turbulence Database (JHTDB) at $Re_\tau = 1000$ (see Appendix B.2), and the full simulation box is displayed in the streamwise and spanwise direction [40].

are alternated by high-speed fluid (for maintaining continuity) and their spacing remains remarkably constant at different Reynolds numbers [16]. Low-speed streaks display a peculiar behaviour known as *bursting*. As streaks move downstream they tend to slowly tilt from the wall, then they rapidly move upwards to finally break into fine-scale motion [4]. Such upward motion from the wall is usually referred as *ejection*. In order to maintain the fluid continuity, ejections are paired with *sweeps*, that is high speed fluid moving downward to the wall. These motions can be effectively characterized, e.g., via quadrant analysis [12, 41, 42] so that ejections belong to the second quadrant (Q2 events) while sweeps belong to the fourth quadrant (Q4 events).
On the other hand, quasi-streamwise vortices are pairs of counter-rotating rolls slightly inclined away from the wall, which sustain and lay in between near-wall velocity-streaks [4, 12]. Quasi-streamwise vortices have a width comparable with that of streaks, but they are less elongated and are responsible for the wall-normal velocity fluctuations in the buffer layer [32]. Furthermore, the presence of quasi-streamwise vortices has been associated with streak bursting [4]. Therefore, since streamwise vortices generated by streaks instability (bursting) tend to reinforce the streak formation, this process lead to a *self-sustaining cycle* [32]. In order to identify such vortical structures and strong shear events, several threshold-based techniques have been employed. In the former case, different methods have been proposed for vortex identification based on the properties of the velocity gradient tensor [43],

while in the latter case variable-interval time (or space) averages (referred to as VITA or VISA, respectively) have been proposed [42].

At higher distances from the wall, in the logarithmic region, streaks are still present but they are much larger and longer than in proximity of the wall [12]. It is worth noting that the physical domain in numerical simulations must be sufficiently large in order to let large structures develop, especially at high Reynolds numbers; a small domain, indeed, could constrain the spatial developments of large coherent structures [12]. The principal vortical structures out from the wall are represented by *hairpin vortices* [39]. These structures are mostly observed in the logarithmic layer as a result of an upwards lifting up of quasi-streamwise vortices, and their size is found to grow with the distance from the wall [14]. It has also been observed that single hairpin vortices tend to cluster in *packets*, which contribute to the transport of vorticity, momentum, and kinetic energy from the wall [39]. Such packets of vortices form the so called large scale motions (LSM), that are observed to extend up to distances of the order of the typical domain size (e.g., boundary layer thickness). As mentioned above, at larger distances from the wall, bigger structures appear as very-large-scale motions whose origin and characterization are still debated [16]. For instance, an open issue remains on the relationship between VLSMs and LSM, as both are associated with large fractions of turbulent kinetic energy [39].

The presence of coherent structures are well captured by correlation functions and spectra. Specifically in wall-bounded turbulence, correlation-based velocity structures are elongated in the streamwise direction and shorter along the transversal directions [44]. Concerning spectra, the energetic contributions due to streaks and (as recently highlighted) due to LSMs and VLSMs are identified as inner and outer peaks in the pre-multiplied spectra of the velocity field [16]. However, it should be noted that the Fourier analysis can only be employed along homogeneous directions and does not take into account location. On the contrary, fixed-point analyses have the advantage of being localized in space but they have no size [32]. An issue that has recently raised attention (thanks to the detection of an outer peak in the pre-multiplied spectra at high Reynolds numbers) is the influence of large- and very-large-scale motions on the near-wall region via a outer-inner interaction [16]. In particular, it has been found that the outer large scales modulate the small-scale behaviour in the near-wall region, namely small- and large-scale fluid motions are coupled (e.g., see Ref.s [45–47] and references therein).

For sake of completeness, we mention other two tools that have been successfully employed for coherent motion characterization, namely proper orthogonal decomposition (POD) [48] and linear stochastic estimation (e.g., see Ref. [49]), whose details are beyond the scope of this work.

Finally, it should be pointed out that, as outlined so far, instantaneous flow fields are usually exploited to investigated coherent structures, and statistical measures are obtained *a posteriori*. In Ref. [50], instead, the authors propose a new concept of *statistical coherent structure*, in which the definition of coherent structure is

directly related to statistical measures rather than an instantaneous field (this view is shared in our work, see Chapter 6).

2.3.2 Turbulent dispersion of passive scalars

In the previous sections, some general insights and open issues on the characterization of wall-turbulence were reported. In this section, we focus on a particular topic that is the dispersion (i.e., mixing and advection) of a passive scalar in a wall-bounded turbulent flow. Turbulent mixing, in fact, is a fundamental issue in many applications from atmospheric and geophysical phenomena, chemical reactions and combustion mechanisms to biophysics [8, 51].
In this work, in particular, we consider two issues: (i) the dispersion of passive scalar through a Lagrangian framework, in which a set of fluid particles is tracked in a numerically simulated (DNS) turbulent channel flow (see results in Chapter 7); (ii) the dispersion of a passive scalar plume emitted from a point source in an experimentally generated turbulent boundary layer (see results in Chapter 5, Section 5.4). Therefore, this section deals with these two issues, by highlighting the main contributions related to such topics.

Dispersion from a Lagrangian viewpoint

Following the seminal papers of Taylor [18, 19] and Richardson [52], the Lagrangian viewpoint on turbulent dispersion has progressively acquired great relevance, as the Lagrangian framework naturally emerges in problems – including modelling – related to (turbulent) transport [53, 54]. Most of the advances in the topic, however, have been carried out in the last few decades, as a consequence of the increased performance in computational and experimental technologies. With respect to the Eulerian viewpoint, indeed, Lagrangian measurements in experiments are more challenging as they requires the intrusion of external particles, which have to be properly tracked with sufficient resolution both in space and time [55].

With the aim of characterising turbulence from a Lagrangian perspective, statistics have been computed in terms of Lagrangian autocorrelations (of velocity and acceleration fields), probability density function (PDF), lower- and higher-order moments, structure functions, as well as one- and multi-particle dispersion statistics [7, 51, 55, 56]. In particular, the importance of particle relative dispersion is due to its relation with mixing and dissipation of the scalar fluctuations, with obvious implications on practical applications [7, 56]. In order to extend the more common two-particle dispersion, the geometrical features of clusters made up of three or four particles have been investigated [57–60], as well as particle-puffs [61]. In fact, Lagrangian statistics of higher-order moments requires that groups of particles are simultaneously tracked [60]. Different from particle pairs, multi-particle dispersion can lead to complex shapes as an effect of turbulence. Accordingly, multi-particle

geometrical features have been usually analysed in terms of volume and size change, where the latter one is typically characterized by the inspection of the moment of inertia of the multi-particle shapes. In particular, studies on isotropic and convective turbulence [57–61] reveal that for long times multi-particle shapes tend to be strongly distorted, specifically by developing into elongated shapes.

However, most of the studies involving passive scalars have focused on homogeneous isotropic turbulence with few applications on relative dispersion on anisotropic and inhomogeneous turbulence [62]. Although turbulence significantly enhances dispersion with respect to molecular diffusion [52], the presence of a mean shear makes the relative dispersion even more complex, by further enhancing the relative separation between particles (especially in the streamwise direction). In fact, while in homogeneous isotropic turbulence statistics depend only on the value of the initial distance between particles, in wall-turbulence statistics depend also on the wall-normal distance and on the direction of initial particle separation [62].

For a turbulent channel flow, there are few works dealing with Lagrangian dispersion of particles. The first numerical studies on Lagrangian statistics were carried out between 1970 and 1990s as a consequence of the availability of numerically extracted data [63]. These investigations were characterized by low Reynolds numbers as well as a relatively low number of particles. More recently, Lagrangian statistics have been acquired at higher Reynolds numbers, also by considering the effects of particle inertia [62, 64–66]. More in detail, in *Pitton et al.* [64], particle separation is investigated in a DNS at $Re_\tau = 150$ for fluid and inertial particles. The authors have found that by removing the effect of the mean shear on particle dynamics, the pairwise separation is much more prevalent in the streamwise direction rather than the wall-normal and spanwise directions. A turbulent channel flow at a higher Re value has been investigated by *Kuerten and Brouwers* [65] in which Lagrangian velocity autocorrelations at $Re_\tau = 950$ were investigated. Moreover, *Stelzenmuller et al.* [66] and *Polanco et al.* [62] investigated channel flows at $Re_\tau = 1440$ via acceleration statistics and (forward and backward) relative dispersion of particle pairs, respectively. In particular, in Ref.s [62, 65] the authors also propose models for the statistics scaling, which are accurate only away from the near wall region.
To conclude, it is worthwhile to mention that similar analyses have also been carried out for other canonical turbulent flows (that are not the object of this work), namely pipe and boundary layer flows (see Ref. [63] for an overview).

Dispersion of passive scalar plumes

Among different configurations that can be explored for the dispersion of pollutants in the atmospheric (turbulent) boundary layer, a typical issue is represented by the release from a point source. This issue, indeed, is related to the dispersion of pollutants, contaminants or toxic substances, with significant environmental and

health impact.

Information on the statistics of the substance concentration is required for characterization and modelling of the dispersion in the atmospheric boundary layer, especially in urban areas. As a result, different strategies have been adopted to investigate the relation between one-point probability density functions (PDFs) of the pollutant concentration and the turbulence dynamics, under different release conditions. With the aim to characterize the concentration PDF, the relation between mean and standard deviation as well as a functional dependency between the third- and fourth-order moments have been investigated by *Chatwin and Sullivan* [67] and *Mole and Clarke* [68], respectively. The shape of the concentration PDF has also been investigated; good agreement is found by modelling the PDFs by means of a Gamma function (e.g., see Ref.s [69, 70]). In general, the typical modelling approaches of concentration PDF involve Lagrangian stochastic models, plume meandering models, micro-mixing models, as well as models based on numerical computation (e.g., large-eddy-simulation) [71]. In particular, meandering models [72–75] are the most simple and rely on the conceptual framework described by *Gifford* [76], in which the dynamics of a plume emitted by a point source is mainly governed by a meandering motion and a relative dispersion of the plume. The meandering is an irregularly oscillating motion of the instantaneous center of mass of the plume; it is due to the effects on the plume dynamics of turbulent length scales larger than the plume size. On the other hand, the relative dispersion refers to the spreading of the plume with respect to its center of mass.

The reliability of these models need to be experimentally verified by means of both concentration and velocity measurements. For a point-source release, the work by *Fackrell and Robins* [77] represented a reference point for several successive studies during the last decades. The authors investigated the effect of the height and size of the source on the concentration (lower-order) statistics in a rough-wall setup. They found that, for an elevated source, plumes emitted by a smaller source size display a stronger meandering motion with respect to plumes emitted by a larger source size. On the other hand, for a ground-level source, the plume dynamics is slightly affected by the source size. *Nironi et al.* [71] successively conducted wind-tunnel experiments by evaluating both lower- and higher-order statistics, thus extending the work by Fackrell and Robins. The authors found that small variations of the source size significantly affect the role of meandering in the plume dispersion, by inducing variations in the (one-point) concentration statistics up to streamwise distances of the order of a hundred times the source size [71, 77]. In particular, while the mean concentration profiles are slightly affected by the source size, the second- and higher-order moments are significantly affected by the emission conditions especially in the proximity of the source [71].

Concerning smooth-wall boundary-layers, only recently *Talluru et al.* [78] investigated the role of streamwise velocity structures on transport of passive scalar plumes, by focusing on the effect of the elevation of the emitting source with fixed

size. For concentration statistics, they found results in accordance with those by *Fackrell and Robins* [77] and *Nironi et al.* [71]; furthermore, they provided statistics of longitudinal turbulent flux which are summarized into a proposed physical model [78]. The same authors have also investigated the scaling of concentration spectra measured at different spatial locations, by founding that a self-similar behaviour emerges if concentration spectra are normalized by concentration variance [79].

In this work, we address the same issue studied by *Fackrell and Robins* [77] and *Nironi et al.* [71] in a rough-wall turbulent boundary layer, by investigating concentration and vertical turbulent transport time-series via classical statistics and complex network metrics (see results in Section 5).

Chapter 3

Complex networks: a novel perspective on complex systems

"Develop your senses, especially learn how to see.
Realize that everything connects to everything else."

Leonardo da Vinci

This Chapter aims to introduce complex networks, by highlighting the historical context in which network science has arisen, the main applications, as well as the mathematical formalism behind network analysis. The Chapter is organized as follows. In Section 3.1, an introduction on complex networks is provided, by focusing on the principal findings that have led to the establishment of network science, and the main difference between a sociology-based and a physics-based network analysis. Section 3.2 comprises a brief summary of the main applications and research fields that have mostly taken advantage from complex networks in last two decades. Finally, complex networks are formally described in Section 3.3. In particular, Section 3.3.1 includes the main definitions, as well as the different way in which a network can be constructed. Quantitative measures are described in Section 3.3.2, by focusing on the set of metrics that have been exploited in this work for network-based analysis of turbulent flows.

3.1 Introduction

Connections are the basis of human activity in almost every context and stage of life. As basic examples, we connect with each other by talking, writing mails, sending messages or through an handshake. Any financial transaction, social collaboration, degree of kinship, sporting event are based on inter-connections. Basically,

we live in an world in which *everything connects to everything else*. Networks arise, therefore, in order to geometrically represent the elements of a (discrete) system and to study the way they mutual inter-connect. However, complexity is not a characteristic of any network, but it typically emerges as real-world systems are analysed. Therefore, complex networks have been usually defined as graphs that show non-trivial (topological or structural) features, and that arise from real-world complex systems.

Historically, social networks were one of the first network applications, developed in 1920s with the aim to understand how individuals connect with each other and what implications such relationships have in the society [80]. It has only been at the end of 1990s that network analysis has started to deeply involve other disciplines. The seminal papers of *Watts and Strogatz* [81] in 1998 and *Barabási and Albert* [82] in 1999 represented the first efforts to highlight the non-trivial texture of interconnections in physical complex systems, showing that real systems tend to display a complex structure of interconnections that deeply deviate from a random or a regular configuration. Since then, complex network analysis has increasingly attracted the attention of scholars from disciplines different from sociology, such as physics and biology. Nevertheless, there are substantial differences in the way network analysis is carried out between the classical sociology-based approach and the novel physics-based one. While social network studies are theory-driven (i.e., theory $\rightarrow$ hypotheses $\rightarrow$ network generation $\rightarrow$ network analysis), while physical network studies are data-driven (i.e., data $\rightarrow$ emergence of patterns in network analysis $\rightarrow$ post-hoc hypotheses or model for pattern) [83]. Figure 3.1 shows a sketch of the processes involved in the network analysis of real-world complex systems.

In last two decades, three key concepts have driven complex network research: (i) the *small-world* effect, (ii) the *scale-free* networks and the idea of *centrality*, and (iii) the occurrence of *communities*. The small-world phenomenon was firstly expressed as the result of a social experiment resulting in the statement that any two people can be reached with each other in at most six intermediate 'friend of a friend' connections. However, *Watts and Strogatz* (unexpectedly) found that also many other real systems display the small world effect [81]. This implies that, in many situations, real systems display stronger interconnected structures than expected.

A second, crucial, idea in complex network analysis is centrality, that is the identification of the most important elements in a network [84]. For instance, in a network in which connections are represented by flight routes between airports, the most central airports could be those with the largest number of departures per day, as well as those with a low number of departures but in a strategic geographical position for intercontinental flights. *Barabási and Albert* found out that many real-world complex systems display a power-law probability distribution (i.e., scale-free) for the number of connections in the networks [82]. This suggested the presence of *hubs* in the networks, namely elements of the systems much more *central* than

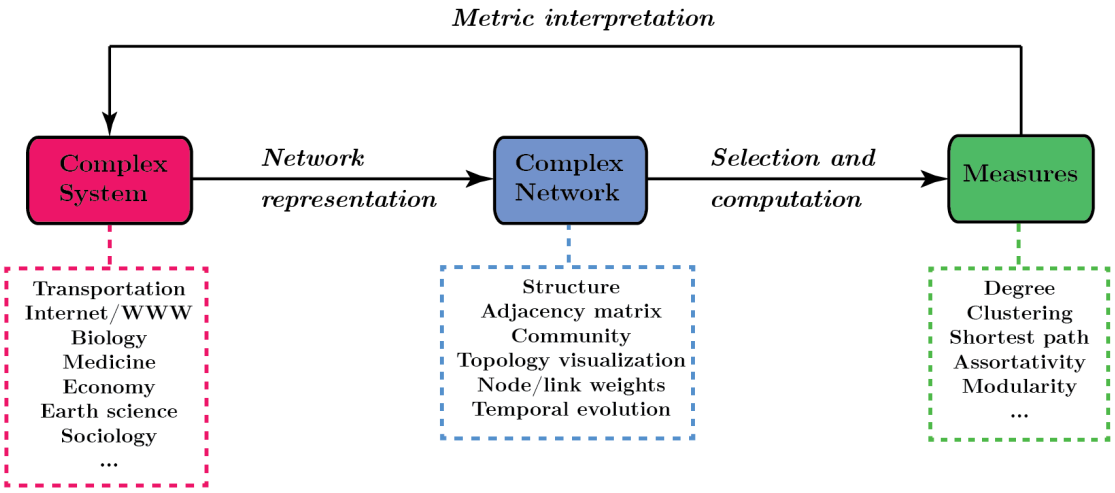


Figure 3.1: Block-diagram showing the typical processes involved in the construction of a physics-based network. The diagram highlights the closed-loop structure of the process, so that measures have to be interpreted in physical terms to extract physical insights.

the remaining elements. Moreover, scale-free networks can emerge as a result of *preferential attachment* mechanism, namely a new element will *prefer* to attach to an existing element of the network that is well connected (i.e., the hub) [82].

Third, the idea of community structure in complex networks has important practical developments. A community is a sub-group (or cluster) of entities that are strongly connected with each other and poorly connected with the rest of the network [85]. As for coherent structures in turbulent flows (see Section 2.3), communities are the manifestation of a self-organization in the system. The tendency to form well-connected sub-structures has a fundamental impact on the modelling and understanding of several processes.

To summarize, thanks to the natural suitability of networks to virtually represent any discrete system – and following the developments of the key ideas of small-world, centrality and community – the last two decades have witnessed a rapid establishment of a cutting-edge approach to research: *network science*.

3.2 Network science: an escalation of applications

The versatility of complex networks in representing discrete systems, in conjunction with the advances in computational capabilities, has led to an extremely rapid development of tools for network analysis in a large variety of applications. The

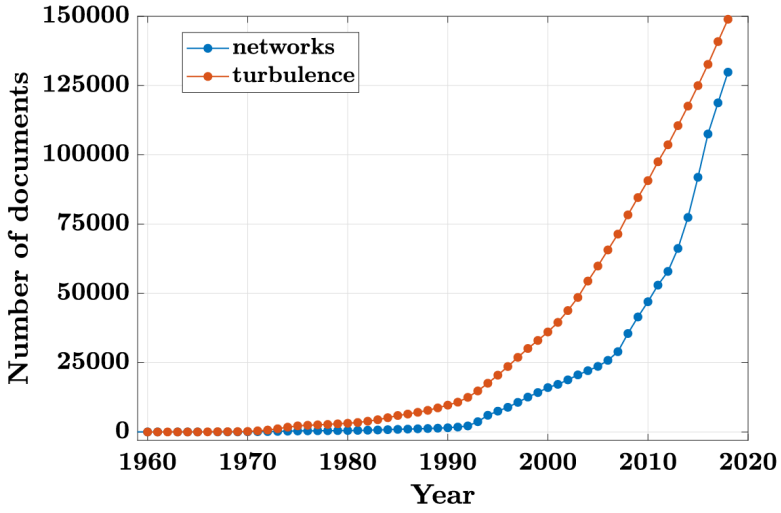


Figure 3.2: Cumulative number of documents from 1960 to 2018 containing the keywords "complex network OR network analysis" (blue line) and "turbulence OR turbulent flows" (red line), as indexed by SCOPUS database.

data-based approach of physical networks has taken great advantage from the enormous production of information and details on real-world systems. To grasp the concept of the huge impact of network science, Figure 3.2 shows a comparison of the (cumulative) number of documents indexed by SCOPUS containing the keywords "complex network" or "turbulence". The number of documents from 1960 to 1990 involved in turbulence is more than six times the number of documents involved in network analysis in the same period. However, the ratio between the number of documents containing keywords "turbulence" or "complex network" from 1960 to 2018 is just 1.15.

Despite the large variety of applications of complex networks to model real-world data, in the last two decades, some research fields have taken more advantage than other from the powerful tools of network science. In the following, a brief survey of the main network-based applications is reported (for details, see Ref. [87]).

- **Biological Networks.** An active area of research in network science is represented by biological networks, which comprises biomolecular analysis, general medicine and neuroscience. Biological processes in living organisms have been investigated via *protein-protein interaction networks*, *metabolic networks* and *genetic networks*. In medicine, network analysis has been exploited for the understanding of diseases and their spreading (e.g., see Figure 3.3), while in neuroscience complex network analyses have been carried out to study brain functions and diseases as well as the effects of environmental factors on the brain formation and evolution.

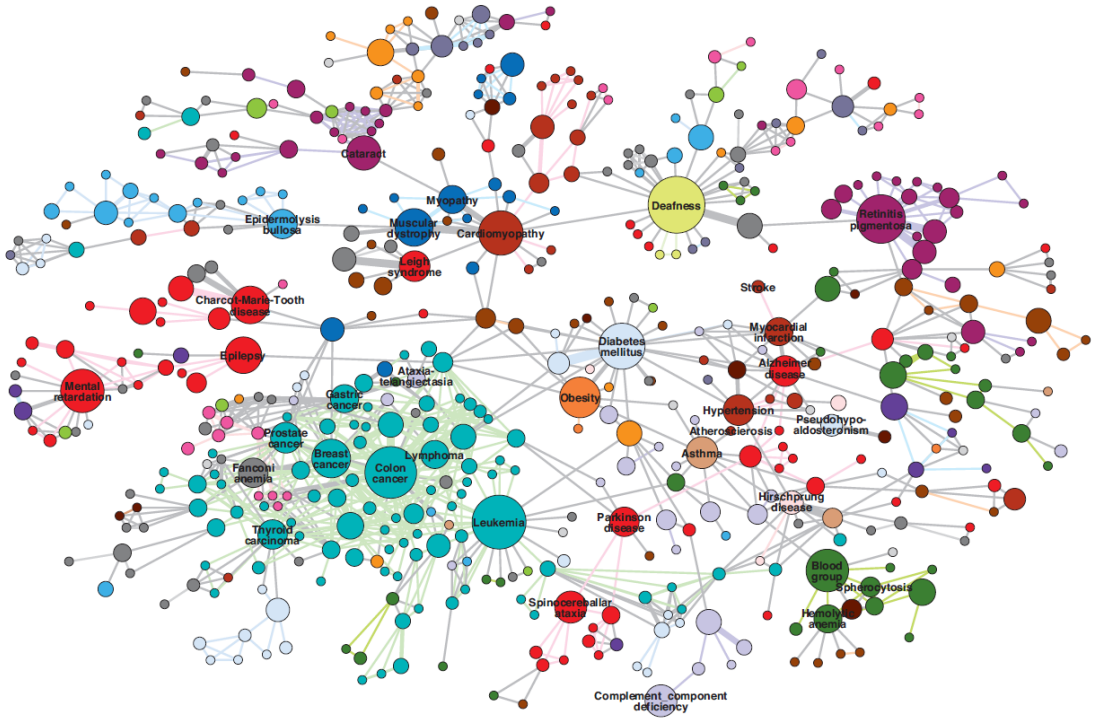


Figure 3.3: Example of humane disease network, where nodes are diseases and links are active if they share one or several disease-associated genes. Node colours indicate the disease class, while node size indicates the number of genes associated with the corresponding disease [86].

- **Social and Economic Networks.** In the last decades, several research fields in social activities – e.g., sport activities, political elections and scientific collaborations – have taken advantage of the network-based approach (e.g., see Figure 3.4). For instance, networks based on scientific collaborations have revealed that the most productive authors are those involved in a great variety of interests. Furthermore, in the context of social networks, the idea of epidemic spreading plays an important role. For example, if an epidemic threshold does not emerge in a social network (i.e. diseases spread regardless of the infection rate), a disease can easily propagate to all nodes of the network with several health implications.

  Economic studies usually concern (among others) the analysis of trades, currency, wealth distribution and tourism. Since economic systems are made up of interacting agents that aim to improve (local or global) economy, complex networks have represented a suitable tool to investigate economic relations. A key point in economy-related networks is the large data availability, especially financial data that have been collected along several years with high accuracy.

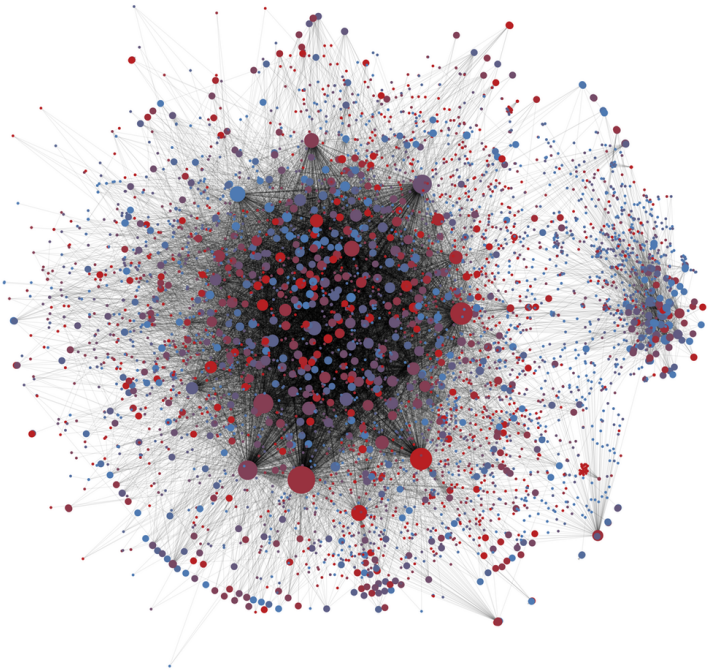


Figure 3.4: Network based on Twitter sharing of the hashtag #SB277 concerning California vaccination law. Nodes are Twitter accounts while links between them show retweeting of hashtagged posts. Larger node size indicates a higher retweet activity. Red nodes are likely bots while blue ones are likely humans [88].

Figure 3.5: Example of airport transportation network, where nodes correspond to airport locations. Colours indicate different geographical areas. [89]

- **Engineering and Physical Networks.** Many phenomena in different areas of engineering and physics can be modelled and investigated through networks. In engineering, transportation systems (e.g., airports and urban

streets) and electric power systems have particularly benefited from network analysis. In the case of airport transportation (e.g., see Figure 3.5), networks provided results for efficient resource allocation as well as for attack vulnerability. Road networks has been studied with the aim to improve accessibility, that is the possibility to easily travel inside cities. Electric power systems have also been studied due to their complexity, as they comprise several transmission lines, substations and generators.

Computer science and the internet (e.g., the World Wide Web) are further examples of research fields involved in network science. In fact, due to their intrinsic web-like structure, data-sharing systems and software architectures have taken great advantage from the network formalism.

Among other physical systems such as particle physics or chemistry, Earth sciences have largely benefited from the complex network approach. Specifically, climate networks represent an active field of research mostly focused on global changes in climate dynamics. For instance, complex networks have been constructed to investigate *El Niño* and *La Niña* events, namely strong growths and decreases in the sea surface temperatures in the central and east-central equatorial Pacific Ocean. More details on climate networks are included in Section 4.3.

- **Other Networks: Linguistics and Ecology.** Finally, it is worth to mention other two applications of complex networks that have important practical implications: linguistic networks and ecology-related networks. In the former case, linguistic networks can promote the study of co-evolution between languages as well as the implementation of user-friendly reliable technologies as search engines and machine translators. In ecology, relationships like competition, parasitism and predator-prey have been investigated with the aim to understand how different species interact with each other, with implications on food-webs.

Among the plethora of applications of complex networks, fluid flows have been also recently investigated via network-based approaches. However, due to the importance related to the present work, the network analysis of fluid flows (especially turbulence) is detailed in the next Chapter 4.

3.3 Definitions and measurements

In this section, formal definitions of a network and the metrics investigated in the present work are reported. Some of the concepts described are typically used in network theory, while some others are here introduced *ad hoc* to better highlight some features of turbulent flows. For a comprehensive review of definitions and measurements, see Ref.s [80, 83–85].

3.3.1 Network formulation

A network is defined as a graph $\mathcal{G} = (\mathcal{N}_v, \mathcal{N}_e)$, made up of a set of $\mathcal{N}_v$ labelled *nodes* (or *vertices*) and a set of $\mathcal{N}_e$ *links* (or *edges*). The structure created by node interactions is called the network *topology*.

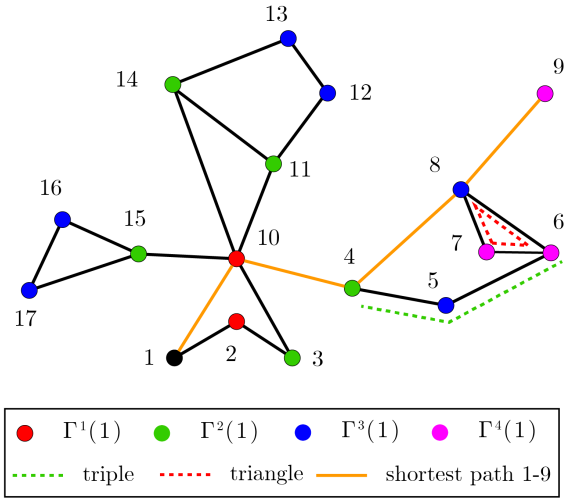
According to the way links are defined, networks can be classified in four categories [85]: (i) *weighted networks*, if a (scalar) real value is assigned to each link, thus quantifying the link intensity; (ii) *unweighted networks*, if link intensity is not taken into account, i.e. links are just present or absent; (iii) *directed networks*, if a link points from a node (source) to another node (target); (iv) *undirected networks*, if link direction is not taken into account (source and target nodes coincide). In this work, network analysis is only limited to undirected networks. Figure 3.6 shows a sketch of a network in which different structural features are highlighted. Specifically, nodes (depicted as coloured dots) can represent time instants, spatial locations or Lagrangian particles in case of time-series, Eulerian or Lagrangian analysis, respectively (see Chapter 4). In order to represent the node-node interconnections, the *adjacency matrix* is defined as

$$\mathcal{A}_{i,j} = \begin{cases} 1, \text{ if nodes } \{i,j\} \text{ are linked, with } i \neq j \\ 0, \text{ otherwise} \end{cases}, \tag{3.1}$$

where $i, j = 1, ..., \mathcal{N}_v$ indicate network nodes. Therefore, the entries $\mathcal{A}_{i,j}$ take into account the presence, $\mathcal{A}_{i,j} = 1$, or absence, $\mathcal{A}_{i,j} = 0$, of a link between each pair of nodes $\{i, j\}$. Accordingly, if $\mathcal{A}_{i,j} = 1$, the nodes $\{i, j\}$ are said to be *adjacent*. The set of nodes directly connected to a generic node i is called the *first neighbourhood*, $\Gamma^1(i)$, of i and the nodes belonging to $\Gamma^1(i)$ are called *first neighbours* of i (e.g., see Figure 3.6). In other words, if $\mathcal{A}_{i,j} = 1$ the node j is in the first neighbourhood of the node i, and *vice versa*. In general, a set of nodes constitutes the n-th neighbourhood, $\Gamma^n(i)$, of a node i if the minimum number of different links connecting i and $\Gamma^n(i)$ is equal to n, with $n \geq 0$ and, by definition, $\Gamma^0(i) = i$ [90]. For instance, the first four neighbourhoods of node $i = 1$ are highlighted in Figure 3.6. The *n-th cumulative neighbourhood*, $\Gamma^n_c(i)$, of a node i is the union of its first n neighbourhoods (including i) [90]. In the example of Figure 3.6, the value of n so that $\Gamma^n_c(1)$ comprises all nodes in the network is $n = 4$, as the whole set of nodes is given by the union of the first four neighbourhoods of $i = 1$.

To take into account link intensities, the *weight matrix*, $\mathcal{W}_{i,j}$, is used in place of the adjacency matrix, where entries $\mathcal{W}_{i,j} \in \mathbb{R}$. Either the adjacency or weight matrix corresponds to a mathematical representation of the network. Since only undirected networks are considered in this work, the adjacency/weight matrices are symmetrical (i.e., $\mathcal{A}_{i,j} = \mathcal{A}_{ji}$ or $\mathcal{W}_{i,j} = \mathcal{W}_{ji}$). It should be noted that an adjacency (binary) matrix can be obtained by applying a threshold-operation to a weight matrix.

Although weights are usually assigned to links to quantify the intensity of an



Network metrics

$\mathcal{N}_v = 17$, $\mathcal{N}_e = 22$
$\overline{\mathcal{K}} = 2.59$, $\varrho = 0.17$
$\mathcal{N}_\triangle = 3$, $\mathcal{N}_\wedge = 45$, $\mathcal{T} = 0.2$
$\overline{\mathcal{C}} = 0.27$
$\mathcal{L} = 2.92$
$r = -0.13$
$\Gamma^1(1) = \{2, 10\}$
$\Gamma^2(1) = \{3, 4, 11, 14, 15\}$
$\Gamma^3(1) = \{5, 8, 12, 13, 16, 17\}$
$\Gamma^4(1) = \{6, 7, 9\}$

Figure 3.6: Sketch of a network (*left*) and some related metrics (*right*). A node i (depicted in black) and its first four neighbourhoods (depicted as red, green, blue and magenta dots, respectively) are shown. An example of shortest path between nodes $i = 1$ and $j = 9$ is illustrated as orange lines, while an example of triple and triangle is highlighted as green and red dashed lines, respectively.

interaction, scalar values can also be assigned to nodes [91]. This technique is useful when nodes represent an heterogeneous set of entities (e.g., see Chapter 6). For example, if nodes indicate cargo ships in a transportation network, a weight proportional to the cargo capacity can be assigned to each node to quantify the possibility of a cargo ship to transport goods in the network. Another example is climate networks, in which nodes correspond to geographical positions resulting into different surface areas at different latitudes. In this last case, a *spatial network* is built, namely a network in which nodes correspond to specific location in space.

Finally, complex networks can be classified into *static* and *dynamic*. A static network is a graph whose elements do not vary in time, while in a dynamic (or *temporal*) network, nodes and/or links change in time [92]. In this work, we consider time-varying networks in which nodes are fixed while activation and weight of links change in time. In this way, it is possible to investigate physical phenomena by highlighting transient behaviours (e.g., see Chapter 7).

3.3.2 Network metrics

Each complex network presents specific topological features that have to be investigated to understand the underlying complex system. Therefore, measurements are essential to characterize, classify and model complex networks [85]. The selection of the *best* network metrics that provide the most useful information on

the physical system under investigation is a non-trivial task. Since, for each complex system, nodes and links can be defined following different assumptions, the choice of a set of metrics depends on the way the network is built and need to be carefully addressed (see Ref. [83] for a compelling review on this topic). In what follows, a report of the metrics exploited for analysing wall-turbulence is provided, by highlighting the definition as well as the main features of each metric. It should be pointed out that, in general, metrics can be associated to single nodes or to the entire network.

Connectivity metrics

One of the simplest outcomes to extract from a network is *how many connections* a node activates. Connectivity metrics, therefore, quantifies to the extent to which a node connects to all the other nodes, thus also representing indexes of centrality. In fact, in general, the more a node is connected the more it is important (i.e., central) in the network. The definition of connectivity depends on network features, e.g. whether the network is weighted or unweighted.

Node degree, $\mathcal{K}$**.** The *degree centrality* of a node is defined as

$$\mathcal{K}(i) \equiv \sum_{j=1}^{N_v} \mathcal{A}_{i,j}, \tag{3.2}$$

thus quantifying the number of nodes linked to i. For example, in Figure 3.6, the degree of node $i = 1$ is $\mathcal{K}(1) = 2$, as there are two (red) nodes linked to $i = 1$. It should be noted that $\mathcal{K}(i)$ is a measure of the cardinality (i.e., number of elements) of the first neighbourhood $\Gamma^1(i)$, namely $\mathcal{K}(i) =\mid \Gamma^1(i) \mid$. The average degree is

$$\overline{\mathcal{K}} = \frac{1}{\mathcal{N}_v} \sum_{i}^{\mathcal{N}_v} \mathcal{K}(i), \tag{3.3}$$

which quantifies the average number of links adjacent to a node in the network.

Network density, ϱ**.** A network metric related to the degree centrality is the *edge density*, ϱ. It is defined as the number of links in the network, $\mathcal{N}_e$, over the total number of possible links, namely

$$\varrho \equiv \frac{\mathcal{N}_e}{\mathcal{N}_v \left(\mathcal{N}_v - 1\right)/2} = \frac{\overline{\mathcal{K}}}{\mathcal{N}_v - 1}, \tag{3.4}$$

where the last equality holds for undirected networks. The edge density is a measure of the sparsity of the network, thus playing a crucial role in the network construction as the computational cost grows as ϱ increases. For example, the network shown

in Figure 3.6 has a density $\varrho = 0.17$, since there are $\mathcal{N}_e = 22$ links over 136 possible connections. In many cases, however, the value of ϱ is very small (e.g., $\varrho \sim 10^{-5} - 10^{-3}$), namely the adjacency matrices of real-world networks are sparse (or can be sparsified via specific assumptions).

Node strength, $\mathcal{S}$. In order to take into account link weights, the *strength* of a node i is defined as

$$\mathcal{S}(i) \equiv \sum_{j=1}^{N_v} \mathcal{W}_{i,j}, \tag{3.5}$$

so that the average strength is $\overline{\mathcal{S}} = \sum_i \mathcal{S}(i)/\mathcal{N}_v$. While the degree, $\mathcal{K}(i)$, quantifies the number of links adjacent to i, $\mathcal{S}(i)$ quantifies the sum of the intensity of the relations between node i and all other nodes [93]. In other words, the strength is the link-weighted variant of the degree.

Node volume-weighted connectivity, $\mathcal{K}^{vw}$. Both the degree and strength do not consider node heterogeneity. To take into account node weights, we define the *volume-weighted connectivity*, $\mathcal{K}^{vw}$, as the node-weighted degree [90], namely

$$\mathcal{K}^{vw}(i) \equiv \sum_{j=1}^{\mathcal{N}_v} \left(\mathcal{A}_{i,j} + \delta_{i,j}\right) \frac{\mathcal{V}_j}{\mathcal{V}_{tot}}, \tag{3.6}$$

in which at each node is assigned a scalar value, $\mathcal{V}_i$, equal to the physical volume that the node represents. Accordingly, $\mathcal{V}_{tot} = \sum_i \mathcal{V}_i$ is the total volume occupied by nodes (i.e., the total volume of the physical domain). In Eq. (3.6), $\delta_{i,j}$ is the Kronecker delta, which is exploited to include self-links in the adjacency matrix (see Eq. (3.1)), thus ensuring that $\mathcal{K}^{vw}(i)$ ranges in the interval (0,1].

Probability distributions. Important information on the behaviour of connectivity metrics is provided by their distribution of probability. For example, as mentioned in Section 3.1, scale-free networks – that are found in many real-world applications – show power-law degree distributions, in contrast to Erdős-Rényi random networks that show Poisson degree distributions [85]. The way the centrality metrics are distributed, therefore, can provide useful information on the structure of complex networks. However, in many real networks, the probability distributions of centrality metrics are heavy-tailed, mainly because of the intrinsic noise due to the finite size network [80]. To overcome this issue, it is usually exploited the cumulative distribution. As a result, we define the cumulative degree, strength and volume-weighted connectivity distributions as $P(\mathcal{K})$, $P(\mathcal{S})$ and $P(\mathcal{K}^{vw})$, respectively. They represent the probability to find a node with degree, strength or volume-weighted connectivity less than or equal to $\mathcal{K}$, $\mathcal{S}$ or $\mathcal{K}^{vw}$, respectively.

Clustering metrics

The connectivity metrics are all based on two-nodes relationships, since they only take into account the direct interaction between a node and all the other nodes. An important feature of many networks is the presence of three-nodes relationships (called *triples*), which are quantified by clustering metrics. More in detail, a triple is a set of three nodes in which (at least) two of them are directly linked to the third node. As a result, a *triangle* is defined a set of three nodes all linked between them [80]. Examples of triple and triangle are illustrated in Figure 3.6 as red and green dashed lines, respectively. Two clustering metrics are usually investigated: the *clustering coefficient* and the *transitivity*.

Clustering coefficient, $\mathcal{C}$**.** For undirected networks, the clustering coefficient is defined as

$$\mathcal{C}(i) \equiv \frac{\mathcal{N}_{\triangle}(i)}{\mathcal{N}_{\wedge}(i)}, \tag{3.7}$$

where $\mathcal{N}_{\triangle}(i)$ and $\mathcal{N}_{\wedge}(i)$ are the number of triangles and triples involving node i, respectively. The clustering coefficient, $\mathcal{C}(i)$, ranges between zero and one and provides information on the extent to which the neighbours of i are connected with each other. For example, in Figure 3.6, $\mathcal{C}(i) = 1$ for nodes $i = \{7, 16, 17\}$, because they are involved in one triangle and one triple, while $\mathcal{C}(i) = 0$ for nodes $i = \{1, 2, 3, 4, 5, 9, 12, 13\}$, because they are not involved in any triangle. Since $\mathcal{C}(i)$ is defined as a node-related metric, the average clustering coefficient can be defined as $\overline{\mathcal{C}} = \sum_i \mathcal{C}(i)/\mathcal{N}_v$. For the example shown in Figure 3.6, $\overline{\mathcal{C}} = 0.27$ that is a relatively small value due to the low number of triangles in the network. High values of clustering coefficient, instead, are typical of small-world networks, as consequence of the well-connectedness between nodes. The formula in Eq. 3.7 can also be generalized for weighted networks by taking into account link or node weights [91, 93].

Transitivity, $\mathcal{T}$**.** Differently from the clustering coefficient, the transitivity is a network-related metric (i.e., transitivity value are not associated to specific nodes) and is defined as

$$\mathcal{T} \equiv \frac{3\mathcal{N}_{\triangle}}{\mathcal{N}_{\wedge}}, \tag{3.8}$$

where $\mathcal{N}_{\triangle}$ and $\mathcal{N}_{\wedge}$ are the total number of triangles and triples in the network, respectively. The factor three in the numerator of Eq. (3.8) accounts for the fact that each triangle comprises three triples, thus ensuring that (as for the clustering coefficient) $0 \leq \mathcal{T} \leq 1$. For instance, in Figure 3.6, $\mathcal{N}_{\triangle} = 3$ (i.e., nodes $\{6, 7, 8\}$, $\{10, 11, 14\}$ and $\{15, 16, 17\}$) while $\mathcal{N}_{\wedge} = 45$, thus resulting a transitivity value $\mathcal{T} = 0.2$. Although the transitivity is conceptually similar to the clustering coefficient, $\overline{\mathcal{C}}$ tends to emphasize the contributions of low-degree vertices, because $\mathcal{N}_{\wedge}(i) \sim$

$\mathcal{K}(i)^2$ that appears in the denominator of Eq. (3.7). In some cases, therefore, the transitivity is exploited instead of the clustering coefficient in order to avoid biases related to low-degree nodes [94].

Assortativity measures

The connectivity metrics reported in Section 3.3.2 provide information on the centrality of nodes in the network. However, connectivity metrics are not directly able to quantify whether the centrality of a node is similar or not to the centrality of other nodes. A network is said to be *assortative* or *disassortative* if nodes tend to link with other similar or dissimilar nodes, respectively, where the similarity is typically measured through a centrality metric (such as the degree or the strength). If nodes do not tend to link neither to similar nor dissimilar other nodes (i.e., there is not a preferred linking tendency), the network is said to be *non-assortative*. In this work, two assortativity measures are considered: the *Pearson assortativity coefficient* and the *average nearest-neighbours connectivity*.

Pearson assortativity coefficient, r**.** The Pearson assortativity coefficient, r, (or simply assortativity coefficient) is ussually defined as the Pearson correlation coefficient of the degree of the nodes at the ends of each link [85]. Since each link can be represented by a pair of nodes $\{i, j\}$, it is possible to assign to each link a pair of node degree values, $\{\mathcal{K}(i), \mathcal{K}(j)\}$. The assortativity coefficient can be evaluated, hence, as

$$r = \frac{\mathrm{cov}\left[\mathcal{K}(i), \mathcal{K}(j)\right]}{\sigma_{\mathcal{K}(i)} \cdot \sigma_{\mathcal{K}(j)}}, \qquad \forall \{i, j\} \;\text{ so that }\; \mathcal{A}_{i,j} = 1, \tag{3.9}$$

where $\mathrm{cov}(\bullet, \bullet)$ and σ represent the covariance and standard deviation, respectively. The main advantage of r, therefore, relies on the ability to quantify network assortativity by means of scalar value. Since it is expressed as a correlation coefficient, r ranges in the interval $[-1, 1]$. Specifically, positive r values are obtained for assortative networks (i.e., if nodes are linked with other nodes of similar degree), while negative or null values of r indicate that the network is disassortative or non-assortative, respectively. For example, in Figure 3.6, the network is slightly disassortative ($r = -0.13$) as high degree nodes (i.e., $\mathcal{K}(8) = 4$ and $\mathcal{K}(10) = 6$) mostly tend to link with low degree nodes. Although the assortativity coefficient is typically defined for the degree centrality, other network metrics can be exploited in Eq. (3.9) (e.g., by replacing $\mathcal{K}$ with the strength).

Average nearest-neighbours connectivity, $\mathcal{K}_{nn}^{vw}$**.** Another way to specify network assortativity is based on the investigation of the relation between the centrality index of a node, i, and the average of the centrality index of the first neighbours of i. In fact, if the nodes linked to i have a centrality index that is

on average similar to the centrality value of i, then the network show an assortative behaviour. In this work, we exploit the volume-weighted connectivity, $\mathcal{K}^{vw}$, to define the average nearest-neighbours connectivity, namely

$$\mathcal{K}^{vw}_{nn}(i) = \frac{1}{\mathcal{K}^{vw}(i)} \sum_{j \in \Gamma^1(i)} \frac{\mathcal{V}_j}{\mathcal{V}_{tot}} \mathcal{K}^{vw}(j), \tag{3.10}$$

thus representing the weighted average of the $\mathcal{K}^{vw}$ values of the first neighbours of i [90]. As for the Eq. (3.6), $\mathcal{V}_j$ is the volume that the node j represents, while $\mathcal{V}_{tot}$ is the total volume occupied by all nodes. If there is no correlation between $\mathcal{K}^{vw}(i)$ and $\mathcal{K}^{vw}_{nn}(i)$ the network is said *non-assortative*; if, instead, $\mathcal{K}^{vw}_{nn}(i)$ is an increasing/decreasing function of $\mathcal{K}^{vw}(i)$ the network is classified as *assortative/disassortative* [90].

Distance-based metrics

According to the assumptions exploited to build a network, different distance-based metrics can be defined. For the fluid-related network analyses performed in this work, three distances are exploited: topological, spatial (Euclidean) and temporal distances. The topological distance between two nodes i and j is expressed as the minimum length of a *path* between i and j, where a path is an alternating sequence of nodes and links (considered only once) between i and j [80]. The *shortest path*, therefore, is defined as the path with minimal cost between two nodes. For example, in Figure 3.6, a shortest path is illustrated as an alternating sequence of orange lines, thus representing the path with the minimum number of links between nodes $i = 1$ and $j = 9$. Spatial and temporal distances are involved whenever nodes correspond to spatial locations or temporal observations. In the former, a spatial network analysis is carried out, while in the latter time-series analysis is performed.

Average path length, $\mathcal{L}$. Following the idea of shortest path, the *average path length*, $\mathcal{L}$, is defined as the average of all shortest path lengths in the network [80], namely

$$\mathcal{L} \equiv \frac{1}{\mathcal{N}_v \left(\mathcal{N}_v - 1\right)} \sum_{i=1}^{\mathcal{N}_v} \sum_{j=1, j \neq i}^{\mathcal{N}_v} d^G_{i,j}, \tag{3.11}$$

where $d^G_{i,j}$ is the length of the shortest path between i and j. If the network is unweighted, the length of the shortest path between two nodes i and j is basically the number of intermediate links between i and j (e.g., in Figure 3.6, $d^G(1, 9) = 4$). On the other hand, if links are weighted, the length of the shortest path, $d^G_{i,j}$, is given by the sum of the link weight over the shortest path.

Weighted physical distance, $\overline{d^w}$**.** In the context of spatial networks, nodes could represent heterogeneous sets of locations (e.g., geographical areas at different latitudes). In this case, the Euclidean distance between nodes have to take into account the node weight. Here, we define the *weighted physical distance* between a node i and a given set of nodes, $\mathcal{R} \subseteq \mathcal{G}$, of the network as

$$\overline{d^w}_\chi(i) \equiv \frac{1}{\mathcal{V}_\mathcal{R}} \sum_j \left(|\chi_i - \chi_j| \, \frac{\mathcal{V}_j}{\mathcal{V}_{tot}} \right), \qquad j \in \mathcal{R}, \tag{3.12}$$

where $\chi \in \{X, Y, Z\}$ are node coordinates, and $\mathcal{V}_\mathcal{R} = \sum_{j \in \mathcal{R}} \mathcal{V}_j / \mathcal{V}_{tot}$ is the (normalized) volume occupied by nodes in the subset of nodes, $\mathcal{R}$ [90]. If node weight is not taken into account, $\mathcal{V}_j = 1$ for any j and $\mathcal{V}_{tot} = \mathcal{N}_v$, so that $\overline{d^w}_\chi(i)$ reduces to the arithmetic mean of the distances $|X_i - X_j|$ (equivalently for Y and Z).

Mean link-length, d^t **and *average peak occurrence,*** ϕ**.** In time-series analysis based on complex networks, a typical approach is to identify the nodes of the network as the time instants of the time-series. For example, if a signal is recorded for one minute at a sampling frequency equal to 100 Hz, the resulting time-series possesses 6000 time instants corresponding to 6000 nodes (one for each time instant). In this case, a temporal distance between nodes can be defined as the difference between the time instants corresponding to nodes. Therefore, the *mean link-length*, $d^t(i)$, of a node i is defined as the mean temporal distance between i and its first neighbours, namely

$$d^t(i) \equiv \frac{1}{\mathcal{K}(i)} \sum_{j=1}^{\mathcal{N}_v} \mathcal{A}_{i,j} \, |t_j - t_i|, \tag{3.13}$$

where t_i and t_j are the time instants associated to nodes i and j, respectively [94]. Since the mean-link length quantifies a temporal distance, it is expressed in time unit of measurement (e.g., seconds). In Figure 3.6, if we identify node labels $1, \ldots, 17$ with time instants (e.g., seconds) we obtain, for instance, $d^t(1) = 5$ because the temporal distances between node $i = 1$ and nodes $j = \{2, 10\}$ are equal to 1 and 9, so that the mean value is 5.

The average of $d^t(i)$ over all network nodes is $\overline{d^t} = \sum_i d^t(i) / \mathcal{N}_v$. In some cases, however, it is useful to highlight low values of mean link-length, thus the *average peak occurrence*, ϕ, is defined as $\phi = 1/\overline{d^t}$ (which is a frequency metric). The name of ϕ is related to the ability of this metric to capture the temporal occurrence of extreme events in the time-series (see Chapter 5) [94, 95].

Chapter 4

Fluid flows and network science: a multidisciplinary approach

"The greatest danger in times of turbulence is not the turbulence. It is to act with yesterday's logic."

Peter F. Drucker

In this Chapter, a review of the principal applications and methodologies developed in the last decade to analyse fluid flows by means of complex network tools is provided. The three main approaches to exploit spatio-temporal data for the network construction are introduced and detailed in the context of turbulent flows. Particular emphasis is given to (i) the current literature to map turbulent flows into networks, and (ii) the procedures adopted in this work, which are introduced as a preparatory material for the result sections.

The Chapter is organized as follows. Section 4.1 presents a brief overview of the approaches to spatio-temporal fluid data, which are examined in detail in the subsequent sections of this chapter. In particular, time-series analysis via complex networks is described in Section 4.2, by focusing on the visibility graph approach. Section 4.3 includes information about the construction of spatial networks from fluid flows in the Eulerian viewpoint, while details on the complex network-based analysis in the Lagrangian viewpoint are reported in Section 4.4.

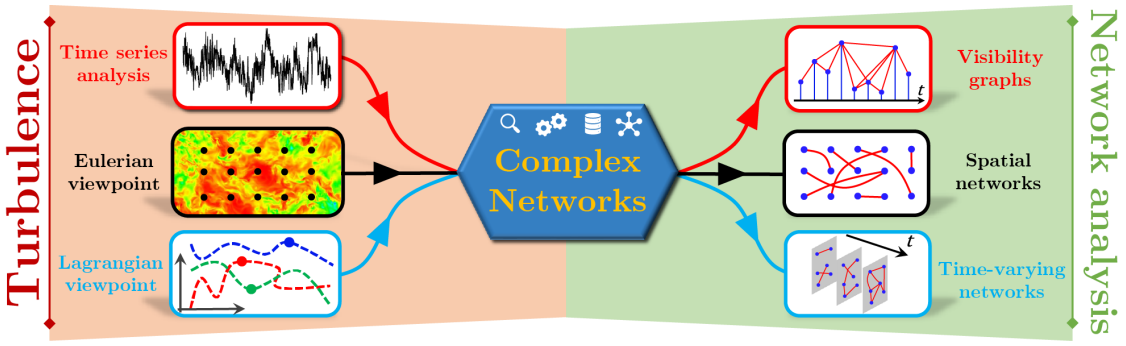


Figure 4.1: Conceptual framework on the different perspectives to analyse spatio-temporal (turbulent) data via complex networks.

4.1 From spatial and temporal data to complex networks

The great progress of network science in developing new tools of analysis (i.e., novel network formulations and metrics) has involved a wide spectrum of disciplines, comprising fluid mechanics. Accordingly, the network-based approach to fluid flows has recently been tackled, for both real-world and prototypical fluid flow applications (e.g., see Ref. [96]). In particular, the approach to data extracted from fluid-related measurements is physics-based, namely spatio-temporal data are exploited with the aim to extract non-trivial features (e.g., recurrent patterns) from the flow dynamics. Following this idea, three main ways to exploit fluid flow data for network analysis have been explored:

1. time-series from non-linear systems are mapped into complex networks, where nodes correspond to (group of) time instants;

2. spatio-temporal fields are mapped into spatial networks, where nodes refer to specific spatial locations in a two- or three-dimensional space;

3. Lagrangian particle trajectories are exploited to build temporal networks (in which particles are associated to nodes) or spatial networks (where particle motion is used as an information medium between different locations).

In this work, these three approaches are carried out with the aim to study wall-bounded turbulence. Specifically, time-series of turbulence quantities are mapped into networks by means of the *visibility algorithm* [97] (see Section 4.2), while the Eulerian approach is based on the Pearson correlation coefficient of spatio-temporal fields, resulting into three-dimensional spatial networks (see Section 4.3). Finally, a Lagrangian viewpoint is adopted to investigate the turbulence mixing effects on the dynamics of a set of fluid particles, which results in a time-varying weighted

network (see Section 4.4). Figure 4.1 illustrates a sketch of the workflow employed to map turbulence data into complex networks.

4.2 A network-based approach to time-series analysis

Time-series analysis – i.e. the study of the temporal evolution of a dynamical system from a recorded series – is a fundamental problem in several research fields. Besides the classical methods (e.g., PDFs, Fourier and wavelet transforms), complex network analysis of time-series has recently emerged as a powerful tool to study non-linear complex systems. The aim is to obtain a graph from each time-series, which is able to inherit and highlight non-trivial information of the system generating the series. With this aim, complex networks have been exploited to study different fluid flows, such as two-phase flows [98–102], turbulent jets [103–106], biomedical flows [107, 108], atmospheric and oceanic flows [109–112], reacting flows [113–117] as well as magneto-hydro-dynamic turbulence [118] and fully developed turbulent flows [119].

From a methodological point of view, three main approaches have been proposed so far to map time-series into networks [120, 121]: (i) *proximity networks*, (ii) *transition networks*, and (iii) *visibility graphs*. In proximity networks, nodes are typically associated with groups of temporal observations called *state vectors*, as they are representative of a state of the system. In other words, time-series are dived into sub-intervals that are exploited as the network nodes. Among the different ways to activate links in proximity networks, correlation- and recurrence-based approaches have been largely employed [121]. Basically, for correlation-based proximity networks, links are active if the correlation between two nodes (representing two states of the system) is sufficiently strong. In recurrence networks, instead, states (i.e., nodes) are embedded in a phase-space (with dimension equal to the length of each state vector) and a link is enabled if two states are sufficiently close in the phase-space. In this last case, distances in phase-space are measured by means, for example, of the Euclidean or maximum norm. Phase-space reconstruction from time-series is a fundamental issue in the context of proximity networks, since it is directly related to the node definition. The typical approach consists in a time-delaying embedding, i.e. to decompose the signal by time-delaying it into sub-series with fixed length (equal to the embedding size), so that each sub-series obtained is embedded in the phase-space. Other approaches can involve signal derivatives for phase-space reconstruction, although they are less common than time-delaying embedding [121].

Another class in complex network analysis of time-series is represented by transition networks. In this case, nodes correspond to discrete states, while a link between two nodes is active if there is a non-zero probability that one node is directly followed

by the other node in time [121]. In other words, links are established if there is a direct transition in time from one state to another state. From a mathematical point of view, this corresponds to a Markov chain with given transition probabilities between discrete states [121]. As a result, transition networks are directed, i.e. links possess a specific direction.
Finally, visibility graphs represent a class of networks that rely on a geometrical criterion to map time-series into complex networks. Differently from proximity and transition networks, in visibility graphs each node correspond to a single value (i.e., a scalar state) of the series. As a result, one of the main advantages of the visibility-based approach is that is non-parametric, while in proximity and transition networks each (vector) state have to be defined (usually involving one to two arbitrary parameters). Links in visibility graphs are established based on a convexity criterion or on a simplified version based on an ordinal criterion. The former is usually referred as *natural visibility graph* (or simply *visibility graph*), while the latter is the *horizontal visibility graph* [121].

It should be noted that, although the three methodological approaches described so far have been usually employed for univariate time-series, each technique can also be extended to multivariate time-series analysis [121]. Accordingly, the multi-dimensional dataset is mapped into multi-dimensional network structures referred to as *multilayer networks*. Each signal in the multivariate set is mapped (by adopting a proximity-, transition- or visibility-based approach) into a graph corresponding to a specific layer in the multilayer network structure. In this way, it is possible to investigate the inter-connections between nodes in the same layer (as in univariate time-series analysis) but also between nodes in different layers, thus unravelling relations between the signals in the multivariate dataset [121].

In this work, thanks to its ability to capture important information (e.g., the presence of extreme events) of the mapped signal, the visibility graph approach is exploited for the univariate time-series analysis of velocity components in a turbulent channel flow (see Section 5.3), and vertical turbulent transport of a passive scalar in a turbulent boundary layer (see Section 5.4).

4.2.1 The visibility graph

The idea to map a time-series into a graph through the visibility algorithm has been proposed by *Lacasa et al.* [97], who have shown that time-series with different structural properties are mapped into networks with different topologies. According to the natural visibility algorithm, two values $(t_i, s(t_i))$ and $(t_j, s(t_j))$ of a univariate time-series $s(t_n)$, $n = \{1,2,...,N\}$, have visibility, and consequently are two connected nodes of the associated network, if the following condition

$$s(t_k) < s(t_j) + (s(t_i) - s(t_j)) \frac{t_j - t_k}{t_j - t_i}, \tag{4.1}$$

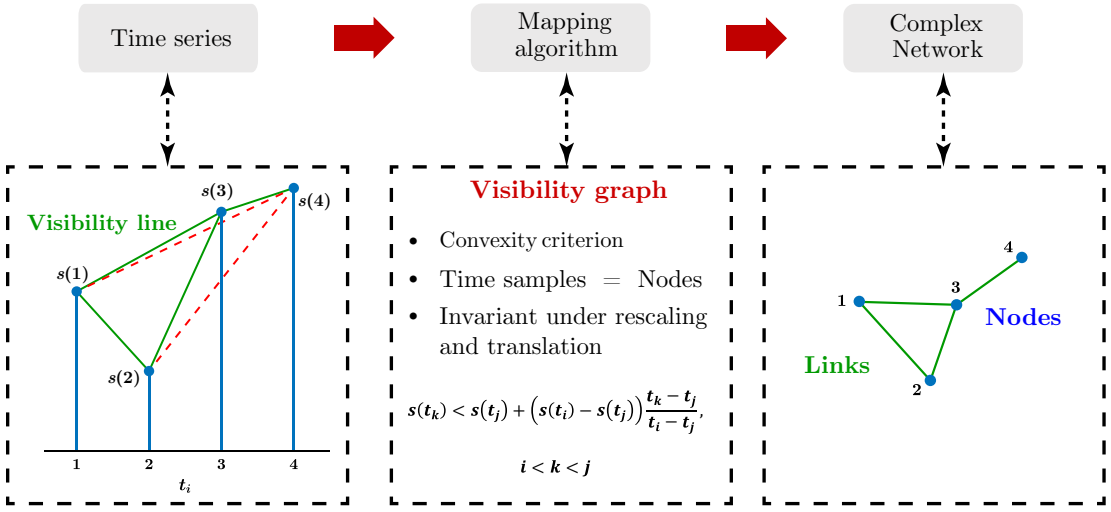


Figure 4.2: Sketch of the process for the building of visibility networks. An example of small time-series, $s(t)$, is displayed as blue stem lines, while the corresponding visibility graph is shown as blue dots and green lines.

is fulfilled for any $t_i < t_k < t_j$ (or equivalently $i < k < j$). Figure 4.2 illustrates an example of the procedure for a small time-series with four values. From a geometrical point of view, two nodes are linked if there is a straight line (called *visibility line*) that connects the two nodes without intersecting any intermediate data. For example, in Figure 4.2, nodes $i = \{1, 4\}$ are not connected because node $j = 3$ obstructs the visibility between them. The condition in Eq. (4.1) is basically a convexity criterion, so that time-series (or sub-intervals) that are convex result into fully connected networks. Furthermore, by exploiting the relation in Eq. (4.1), the visibility networks are undirected and unweighted.

The main advantage of the visibility algorithm is that it does not require any *a priori* parameter. However, the visibility graph is invariant under rescaling and translation of both horizontal and vertical axes (i.e., affine transformations) [97]. The invariance to affine transformations implies that two time-series with different mean and standard deviation values but with similar temporal structure, are mapped into the same visibility network. If the analysis should be sensitive to affine transformations of the series, the invariance represents a drawback of the method. On the other hand, if the focus is on the temporal structure of the series (as in the present work), the invariance is a potential benefit. In fact, it is possible to exploit the visibility graph approach to analyse series without a pre-processed normalization of the series.

The visibility graph approach has been widely employed and different variants of the algorithm have been proposed. The horizontal visibility graph is the main variant of the (natural) visibility graph, and it can be obtained considering only

horizontal lines among data [122]. Differently from the natural visibility graph, the horizontal visibility algorithm satisfies the ordering criterion

$$s(t_k) < \min\left[s(t_i), s(t_j)\right], \qquad \text{for any } t_i < t_k < t_j. \tag{4.2}$$

Due to its simplicity, exact results for the horizontal visibility applied to random time-series have been obtained [122]. Moreover, differently from the natural visibility algorithm, the degree distribution of horizontal visibility graphs is independent from the PDF of the mapped time-series [123]. Accordingly, the degree distribution in horizontal visibility graphs can be thought as a direct measure of the (linear and non-linear) temporal dependence existing in the time-series [122]. A review of the other minor variants of the visibility graph is reported in Ref.s [121, 124].

From a computational point of view, a naive visibility algorithm requires $O(\mathcal{N}_v^2)$ operations to map $\mathcal{N}_v$ data of a series in a network (see also Appendix A). To computationally improve this procedure, optimized approaches have been proposed so far [125–127]. A fast ®MATLAB implementation of the natural and horizontal visibility algorithms has been developed, which relies on a divide-and-conquer approach [125] (the codes are available online, see Ref.s [128, 129]).

4.2.2 Visibility applications to turbulent flows

In the context of turbulent flow analysis, both the natural and the horizontal visibility have been exploited. In 2010 *Liu et al.* [119] took advantage from the visibility algorithm to study time-series of energy dissipation rates in a three-dimensional fully developed turbulence. The degree distribution of visibility networks built from the energy dissipation rates were found to follow a power-law behaviour in the distribution tail. This suggested that the presence of self-similarity in the energy dissipation rate signals. Additionally, the authors evaluated the assortativity coefficient (Eq. (3.9)), which was found to be positive (assortative network).
Few years later, *Charakopoulos et al.* [104] employed the visibility algorithm to discriminate different regions in an experimental turbulent heated jet. By investigating various network metrics at different spatial locations – such as the clustering coefficient (Eq. (3.7)) or the average path length (Eq. (3.11)) –, the authors were able to identify the effects of small and large vortical structures on the time-series of temperature. They were also able to locate the jet axis position as corresponding to the local minimum value of the network metrics.
Turbulent jets have further been studied via the natural and horizontal visibility algorithms. In 2015, *Manshour et al.* [105] applied the horizontal visibility algorithm to acceleration series in a low-temperature helium turbulence, extracted from hot-wire measurements. As in Ref. [119], the authors focused on the evaluation of the degree distribution and network assortativity. By exploiting data at different Reynolds numbers, a stretched exponential function was proposed to model the behaviour of degree distributions, with slope decreasing as Re increases. Moreover,

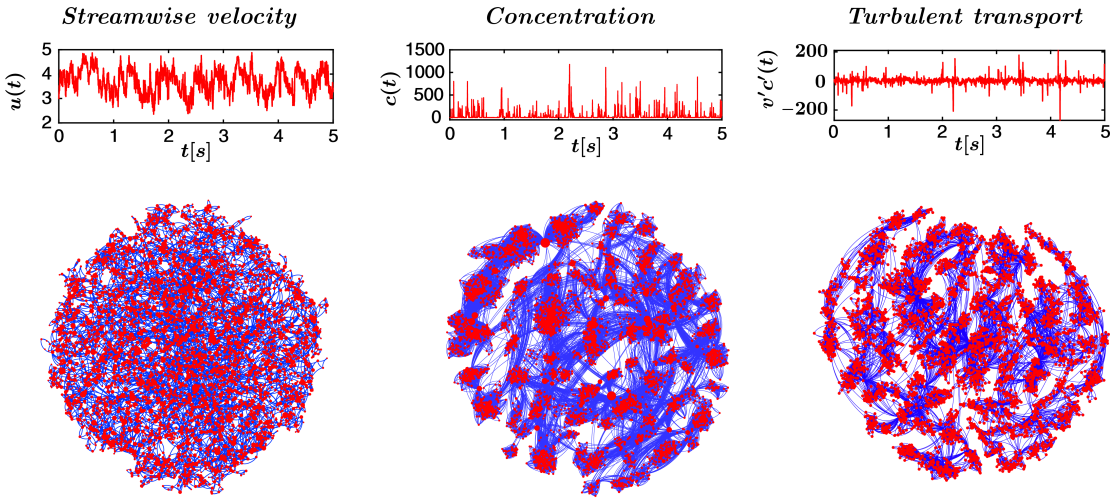


Figure 4.3: Examples of time-series (top) and the corresponding network topology (bottom). Network visualizations are obtained through the *OpenOrd* layout algorithm [130]. The streamwise velocity (left), passive scalar concentration (middle), and vertical turbulent transport (right) are considered. Data are experimentally measured in a rough-wall turbulent boundary layer [95]. Time-series are shown in (m/s), (ppm) and (ppm $\cdot$ m/s) from left to right, respectively.

the visibility networks were found to be assortative, with increasing values of assortativity coefficients as the Reynolds number increases. Recently, *Murugesan et al.* [106] investigated the forced synchronization of a hydrodynamically self-excited jet, by mapping time-series of streamwise velocity intensity via natural visibility. The authors explored several metrics (namely the degree, clustering coefficient, average path length and assortativity coefficient) to study the route to synchronization, by finding that metric values change for different routes to synchronization. The visibility algorithm has also been used in the context of turbulent reacting flows. By mapping time-series of unsteady pressure measured in a combustor, *Murugesan and Sujith* [113] found that acoustic pressure in the combustor can be represented as a scale-free network. In particular, the scale-free behaviour of combustion noise is due to the presence of turbulence. By exploiting network metrics, therefore, visibility networks can provide early warning for the onset of combustion instability.

In Chapter 5, results from the application of natural visibility algorithm to wall-bounded turbulent flows are reported. Differently from the previous works, a specific emphasis is given to the characterization of the temporal structure of the time-series. For example, Figure 4.3 shows three different (experimentally measured) signals and the corresponding network topology. As can be seen, different time-series are mapped into different visibility networks that, hence, are able to

inherit the series temporal structure. Consequently, network metrics are exploited to quantitatively infer the different structure of the signals in terms of extreme events (i.e., peaks in the series) and small variations. A physical interpretation of the metric behaviour along spatial directions is provided, thus linking the temporal structure of the time-series with the underlying turbulence dynamics.

4.3 Spatial networks in the Eulerian viewpoint

From a practical point of view, physical systems are typically represented through sets of data that are discrete in time and/or in space. This is the case of turbulent flows in which, for example, data are acquired numerically or experimentally at specific grid points or spatial locations, respectively. When nodes represent specific spatial positions, the corresponding graph is named as *spatial network*. Thanks to the ability to explicitly represent a discrete set of elements, therefore, complex networks provide a suitable tool for the analysis of discrete systems embedded in a two- or three-dimensional space.

4.3.1 Methodologies and Applications

A plethora of possible spatial networks can be defined from real-world systems (e.g., transportation networks, internet, power-grids among others), according to the node definition. For fluid flows, the principal application – also from an historical point of view – is represented by climate networks (see, among many others, Ref.s [131–140]); other important applications include geophysical flows [141–146], atmospheric dispersion [147], isotropic turbulence [148], vortical flows [149, 150] as well as hemodynamics [151] and thermoacoustic combustion [152–157].

In climate networks, nodes correspond to regions of the Earth surface while time-series of climatological variables (e.g., temperature, rainfall, pressure, or similar quantities) are exploited to activate links [158]. The Earth surface is typically represented by a uniform grid, where each region of the grid is associated to a network node. A uniform grid implies that nodes represent different surface areas at different latitudes. In fact, grid areas close to the equator are larger than close to the poles. As a result, in order to take into account that nodes represent an heterogeneous set, a weight proportional to the surface area is typically assigned to nodes in climate networks [158]. Links are usually activated based on the Pearson correlation coefficient (which accounts for linear dependencies) or the mutual information (which also accounts for non-linear dependencies). Despite the climate network approaches are based on the zero-lag correlation coefficient, weighted networks based on the maximal and minimal cross-correlation at different time-lags have also been investigated [158]. By doing so, complex networks are able to reveal non-trivial patterns in the atmospheric circulation, such as *teleconnections*,

via different climatological variables. Teleconnections refer to climate relations between geographically remote regions, farther than the correlation length scale of the variable. In particular, climate teleconnections are mainly caused by the energy transport and propagation of waves, providing information about the recurrence of climate variability of distant locations.

Recently, the concept of *flow networks* have been pursued for geophysical studies of transport phenomena. As in climate networks, nodes correspond to geographical locations. However, links are not based on correlations but via tracking of substances between different nodes [158]. For example, when fluid transport is focused, fluid particle trajectories can be exploited to connect different regions. By doing so, a link is established if an exchange of fluid occurs between two nodes in a given time interval, and the weight of the link is proportional to the amount of fluid transported [158]. The relation between correlation-based climate networks and flow networks was investigated in Ref. [142]. Since Lagrangian trajectories are exploited to activate links, this approach is also referred as *Lagrangian flow network* [143]. However, it is worth to underline that – despite the terminology – flow networks are actually spatial networks, since nodes represent fixed geographical locations.

Krueger et al. [150] recently took advantage from network formulation to provide a method that is able to compare and classify different vortical flow fields. With this aim, a spatial proximity criterion was employed to activate links between nodes, which represent critical points of the velocity field in a von Kármán vortex street setup. Concerning vortical flows, *Taira et al.* [149] proposed a network formulation based on the Biot-Savart law with the aim to investigate vortical flows in a 2D decaying turbulence. More in detail, a discretization of the physical domain in square cells was carried out, so that each network node corresponds to a grid cell. Since a (vortical) fluid element in a given position induces a velocity in another position (following the Biot-Savart law), the interaction between two locations can be quantified via the amount of induced velocity. Therefore, links were activated for each pair of nodes, and a weight was assigned to each link equal to the average induced velocity between the two nodes. By doing so, the authors found that the strength probability distribution (see Section 3.3.2) follows a power-law behaviour, whose exponent change as turbulence decay.

Following the approach by *Taira et al.* [149] relying on vorticity-induced velocity interactions, *Murayama et al.* [153], *Takagi et al.* [154] and *Hashimoto et al.* [156] analysed the dynamics of the velocity and vorticity fields in a turbulent combustor during the thermoacoustic combustion oscillations. This phenomenon has important practical applications (such as gas-turbine power plants and propulsion systems in aircraft and rocket engines), since strong vibrations induced by large-amplitude thermoacoustic instability can lead to severe structural damages. Moreover, the velocity and the acoustic power fields have also been exploited to

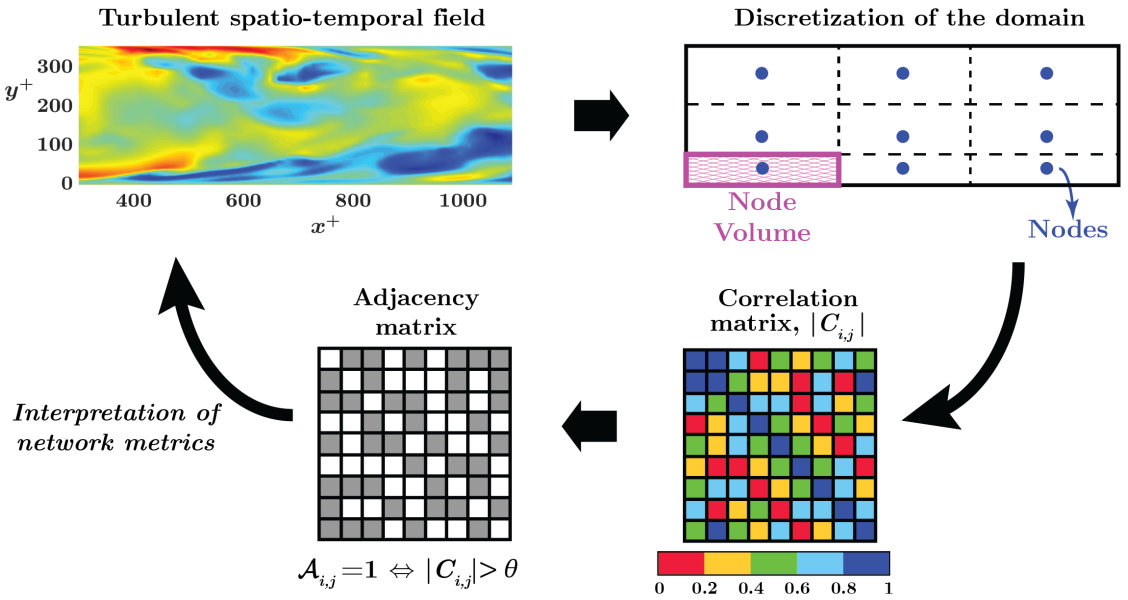


Figure 4.4: Sketch of the construction of a correlation-based spatial network from a turbulent spatio-temporal field. The top-left image shows the streamwise velocity fluctuations, u', in a 2D section of a turbulent channel flow, while in the adjacency matrix, ones (i.e., links) are depicted in grey.

build spatial networks based on the correlation [152] and a spatial-proximity criterion [157], respectively. In both works, the authors investigated the spatial flow dynamics at the onset of thermoacoustic instability, which is a phenomenon that requires to be mitigated in a turbulent combustor. By exploiting the network-based approach, straightforward strategies for mitigating oscillatory instability can then be pursued, as shown in Ref. [159].

Although turbulent spatio-temporal fields have been considered as case-study, spatial network analyses have been basically limited to two-dimensional flows. The first three-dimensional analysis of turbulence was carried out by *Scarsoglio et al.* [148], who built a spatial network by exploiting the correlation coefficient of the turbulent kinetic energy in homogeneous isotropic turbulence. The authors found out that, despite of the isotropy and homogeneity of the flow, the network is able to reveal spatial coherence associated to vortical structures. In order to extend the approach pursued in Ref. [148] to wall-turbulence (which is largely involved in practical applications), spatial networks obtained from three-dimensional turbulent fields in a channel flow are focused in this work.

4.3.2 Spatial networks from 3D turbulent flow fields

By following the ideas developed so far about spatial network analysis of fluid flows (especially of climate networks), in this work a numerically simulated turbulent channel flow is investigated. An example of the procedure carried out in this work is shown in Figure 4.4. The nodes of the network are associated to grid points (i.e., spatial locations) in the domain. Due to the inhomogeneity of the flow in the wall-normal direction, a non-uniform spatial discretization in the vertical direction is adopted so that a weight is assigned to each node which indicates the spatial extension of that node, namely the volume of the corresponding grid cell. As a consequence, nodes at different wall-normal coordinates have different weights.

To activate links, time-series of velocity are exploited in order to evaluate the correlation coefficient, defined here as

$$C_{i,j} = \frac{\mathrm{cov}\left[u\left(X_i, Y_i, Z_i; t\right), u\left(X_j, Y_j, Z_j; t\right)\right]}{\sigma_{u_i} \cdot \sigma_{u_j}}, \tag{4.3}$$

where u is the streamwise velocity, (X_i, Y_i, Z_i) are the coordinates of node i and σ_{u_i} is the standard deviation of u in point i (the same holds for j). The streamwise velocity component is used in Eq. (4.3) for the definition of $C_{i,j}$ as u is the principal variable investigated in this work. However, results for the wall-normal velocity components are also reported and compared to the ones for the streamwise component (see Chapter 6). As illustrated in Figure 4.4, only the highest values of correlation coefficient are retained, by thresholding $C_{i,j}$. Therefore, a link in the network is established if $|C_{i,j}| > \theta$, where $\theta \in (0,1)$ is a threshold value. In general, as for climate analysis, a high value of the threshold θ is chosen to highlight the strongest correlations.

Due to its simplicity and its broad use in the turbulence literature [160] and spatial networks [158], the correlation coefficient represents the most suitable metric to start showing the potential of complex networks applied to wall-bounded turbulence. It is worth to stress that $C_{i,j}$ explicitly depends on the three coordinates of nodes i and j. In turbulence analysis, instead, the average correlation coefficient, $\langle C \rangle$, is typically considered, which (differently from $C_{i,j}$) does depend only on the wall-normal direction, as the operator $\langle \bullet \rangle$ averages in the homogeneous directions and time. Therefore, through the network formalism – namely by exploiting $C_{i,j}$ – the spatial information of the correlation coefficient is preserved by the active links in all directions.

Finally, it should be noted that spatial networks are usually very large (i.e., they possess a huge number of nodes and links), since a high number of grid points is usually required to sufficiently resolve the flow field. In terms of computational cost (see Appendix A), the evaluation and thresholding of all the correlation coefficient values, $C_{i,j}$, is basically the same as that for evaluating the average value, $\langle C \rangle$. However, the size of the resulting correlation-filtered matrix is usually very large

and, although the memory cost is limited due to network sparsity (as discussed in Section 6.2), the implementation of post-processing procedures to extract network metrics could require, in some cases, optimized approaches that deal with a large number of nodes and links. Accordingly, this aspect should be carefully managed in three-dimensional flows, by properly selecting the grid spatial resolution and the value of the threshold, θ. The former, in fact, directly affects the number of nodes in the network, while the latter affects the number of links (since the number of high correlation coefficient values decreases if θ increases). In order to tackle this issue, different approaches have been developed. *Nair et al.* [161] prosed a network sparsification based on spectral properties of the graph and tested the method to a vortical flow field. In Ref. [162], instead, *Meena and Taira* developed a network community-based approach to formulate reduced-order models; the technique was applied to the dynamics of vortical flows and to wake flows over a NACA 0012 airfoil. Recently, *Bai et al.* [163] suggested to reduce the size of the adjacency matrix by applying a randomization technique based on a sampling of the matrix rows and columns. As for the previous mentioned studies, the authors exploited a vortical flow to test their approach. In this work, the value of the threshold, θ, is selected as a trade-off between a high number of nodes (due to a too low θ value) and the link significance (due to a too high θ value that mainly enables strong but trivial connections). Moreover, to decrease the network size in terms of nodes, the spatial discretization of the domain in the homogeneous directions is increased with respect to the numerical simulation. Specifically, the spatial discretization is selected in order to be smaller than the (average) decorrelation length scales in the homogeneous directions.

The results of the spatial network analysis of the turbulent channel flow are reported in Chapter 6. The analysis is carried out at different network levels, namely a global level (to explore the features of the entire network), a mesoscale level (to highlight the features of groups of nodes), and a local level (to study the relation between representative nodes and the remaining nodes in the network). A parametric analysis of the results is also provided to characterize the effects of different correlation threshold values on the network structure. In this way, a characterization of the turbulent channel flow is given via a spatial network, which is able to retain – through the texture created by links – the spatial information of high-correlation values between different locations in the domain.

4.4 Complex networks from Lagrangian particle trajectories

Besides the Eulerian approach that has been mainly fostered by climate studies, the Lagrangian viewpoint has been recently adopted for complex network analysis of fluid flows. Differently from spatial networks where nodes correspond to fixed

spatial locations, in Lagrangian networks nodes are associated to particles that move (passively or actively) in the fluid domain. Since there is a univocal relation between a particle and its trajectory, a node can represent either a particle or its trajectory. In the former case, an attribute is assigned to each node equal to the particle position (which changes in time and in space). We recall that in geophysical spatial networks (also referred as Lagrangian flow networks) particle trajectories are exploited as a mean that conveys an information about fluid transport, but nodes are fixed spatial locations. In the Lagrangian viewpoint, instead, nodes do directly represent particles (or their trajectory).

According to the way in which links are defined, two main categories of Lagrangian networks have been proposed: *proximity-based* and *similarity-based* networks. In proximity-based networks, a link is established if two particles come sufficiently close in space during their motion; on the other hand, in similarity-based networks, two particles are linked with each other if they follow similar trajectories. By exploiting these two concepts, complex network tools have been employed to perform spectral graph analysis for coherent structures identification [164–166] and particle clustering [167, 168], estimation of parameters [169], as well as mixing analysis [170]. In what follows, details on the techniques and applications of the network formulation in the Lagrangian viewpoint are reported (see Section 4.4.1) and the methodology employed in this work is introduced (see Section 4.4.2).

4.4.1 Methodologies and applications

A proximity-based approach was firstly carried out by *Hadjighasem et al.* [164], who exploited spectral graph properties for the coherent vortex detection. By following the principle that coherent structures can be characterized by a set of particles displaying short relative distances over time, the authors tested their approach to three cases: a Bickley jet flow (which serves as a model for zonal jets in the atmosphere), an ocean surface dataset, and an Arnold-Beltrami-Childress flow. Weighted graphs were built, where nodes were associated to particle trajectories and the link weight is inversely proportional to the trajectories relative distance. In other words, a time-averaged Euclidean distance between trajectories is exploited to weigh links, so that closer trajectories result in higher link weights [164]. By doing so, the authors were able to identify coherent vortices in a robust way, even in absence of a preliminary knowledge of the number of present vortices in the domain.

Following the idea pursued in Ref. [164], other proximity-based works have been carried out to study turbulent superstructures in Rayleigh-Bénard convection [167] and mixing in geophysical flows [168, 170]. Specifically, a partitioning of the flow based on spectral properties of the network was performed by *Schneide et al.* [167], in order to detect very-large coherent structures in a numerical setup of a Rayleigh-Bénard convection with large aspect ratio. The authors found that

the resulting clusters match well with the patterns created by temperature superstructures. In the context of geophysical flows, spectral graph partitioning was employed by *Chakraborty et al.* [168] who investigated the presence of dense particle subclusters in coherent vortices, thus highlighting the regions of strong mixing. Finally, *Padberg-Gehle and Schneide* [170] – differently from the other approaches – exploited classical network metrics (such as degree centrality and clustering coefficient) to investigate Lagrangian transport and mixing. The authors (by following the idea of *encounter volume* by *Rypina and Pratt* [171]), focused on two case studies, a Bickley jet and a stratospheric polar vortex. They found that network metrics can be effectively used as indices to quantify the intensity of spatial mixing.

In 2017 *Schlueter-Kuck and Dabiri* [165] proposed an alternative way – called *coherent structure colouring* – to weigh links between particles, which is based on a criterion of similarity between particle trajectories. The authors exploited the trajectory information (i.e., particle positions at each time) to calculate the standard deviation of the relative distance between two fluid particles, normalized by the average distance during the same period. The link weights were then associated to the normalized standard deviation between trajectories. In this way, since the standard deviation can be viewed as an index of (dis)similarity between two datasets, each link quantifies the extent to which two trajectories are similar over a given time interval. The authors exploited the graph spectral properties of the network to identify coherent structures in the flow, by testing their approach to a quadruple gyre flow (which serves as a model for oceanic flows), a Bickley jet and a vortex fluid ring ejected from an axisymmetric nozzle. In a successive work, the same authors employed their coherent-structure-colouring technique to estimate model parameters with potential implications in the context of geophysical flows [169].

Finally, based on the similarity-based approach, *Husic et al.* [172] extended the idea pursued in Ref. [165] by providing a *simultaneous coherent structure colouring* methodology. While in the classical coherent-structure-colouring technique the eigenvector associated to the largest eigenvalue of the graph Laplacian was exploited to detect coherent vortices [165], *Husic et al.* used all the eigenvectors to simultaneously cluster the system states. As a result, the authors were able to provide information on the number, size and shape of vortical clusters, which naturally emerge without prior knowledge of the coherent structure features.

4.4.2 Temporal network-based analysis of turbulent mixing

Although the Lagrangian viewpoint has been adopted for the network analysis of fluid flows, most of the works address two-dimensional geophysical flows (as discussed in Section 4.4.1). Additionally, Lagrangian particle trajectories have been only exploited to build static networks, namely the temporal information intrinsic to each trajectory is enclosed in each link. In other words, each link has usually been assumed to represent a relation between two particles over a time interval.

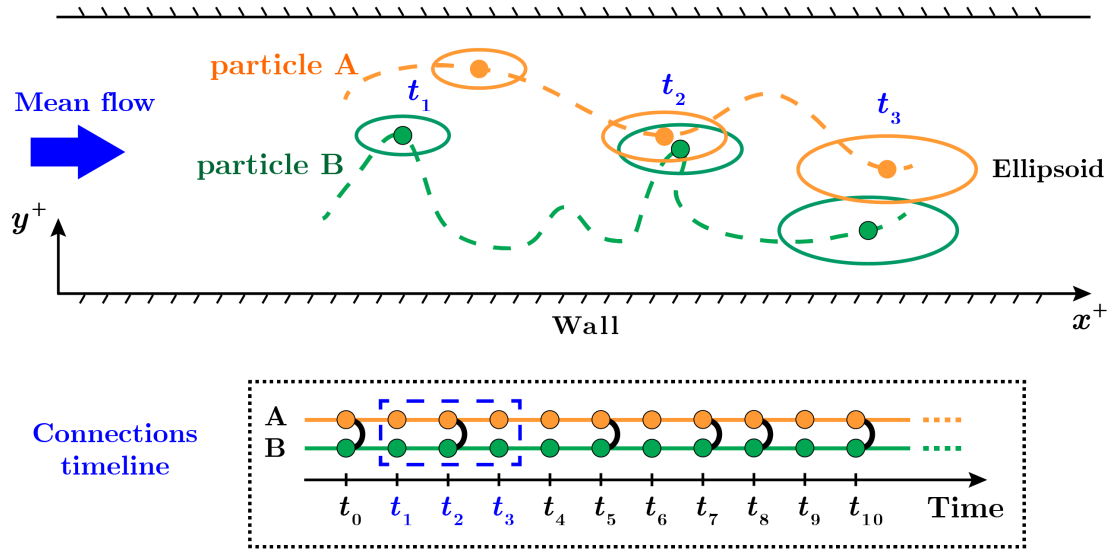


Figure 4.5: (Top) Sketch of the setup for the Lagrangian-based analysis, comprising two trajectories (dashed orange and green lines) of two particles A and B (orange and green dots) and their reference ellipsoids. (Bottom) Time-sequence of the connections between particles A and B (the corresponding sampling time is highlighted by a blue dashed box).

In this work, a three-dimensional setup of a numerically solved turbulent channel flow is employed to study turbulent mixing by means of a set of fluid particles (i.e., tracers). Differently from previous studies, links are here established at each time-step, thus resulting in a time-varying network. Specifically, a time-varying network is a time-ordered sequence of single networks built at each time, and modelled as an ordered set of adjacency or weight matrices [92]. In order to activate links, a spatial-proximity criterion is exploited, namely if two particles come sufficiently close in space during their motion a connection is established between them. For example, Figure 4.5 shows a sketch of the setup, in which two particles move in the domain, and the corresponding connection timeline, which highlights when particles connect with others over time.

To specify the particle proximity, we assume that a particle i is connected to a particle j if i lies inside a reference ellipsoid centred at j, and *vice versa* (by symmetry). In the example provided in Figure 4.5, the two particles are connected at the time t_2 in the time-window considered, since particle A is inside the ellipsoid of particle B (and *vice versa*). The ellipsoid is geometrically anchored to each particle location and it is chosen as reference geometry to take into account the anisotropy of the flow. Therefore, while in previous works a spherical reference geometry was adopted resulting in one parameter equal to the sphere radius [164, 167, 168, 170], three geometrical parameters are required in our formulation, corresponding to the

semi-axis lengths of each ellipsoid.
The choice of the semi-axis length (as the choice of the sphere radius in previous works) is generally a non-trivial task which depends on the specific problem under study. The ellipsoid semi-axes are here indicated as $\{a_x, a_y, a_z\}$ and set proportional to the average pairwise distances between particles in each Cartesian directions. In this way, the increase of the average mutual distance between particles with time – that is mainly due to the streamwise dispersion and partly to the spanwise mixing – is taken into account. In more detail, $a_\chi(t) = \alpha_\chi \cdot \overline{d_\chi(t)}$, where $\overline{d_\chi(t)}$ is the average Euclidean distance between all particles along the direction $\chi = \{x, y, z\}$, while α is a proportionality constant. Although α may also explicitly depend on the Cartesian direction and time, it is set for simplicity as a constant value in the range $\alpha \in (0,1]$. In so doing, the ellipsoid size in each Cartesian direction does not exceed the average pairwise particle distance (this is crucial along the wall-normal direction since particles can not exceed the inter-wall height).

Finally, it should be noted that the approach proposed by *Padberg-Gehle and Schneide* [170] for the study of Lagrangian transport and mixing in geophysical flows was obtained by checking whether a pair of particles comes sufficiently close in space at least once in the time window considered (so that the temporal details of particle trajectories do not explicitly emerge). In this work, instead, the flow anisotropy and temporal dependence are explicitly taken into account. The results of the application of the time-varying network formulation are reported in Chapter 7. Specifically, since the spatio-temporal evolution of particles is captured at each time step, a rich and detailed time-dependent picture of turbulent mixing is provided through the analysis of several network metrics, thus highlighting the temporal development of particle dynamics due to the turbulent motion.

Chapter 5

Visibility-based analysis of time-series in wall-turbulence

Some of the contents presented in this Chapter have been previously published. Below are provided the references:

G. Iacobello, S. Scarsoglio, and L. Ridolfi. "Visibility graph analysis of wall turbulence time-series". In: *Physics Letters A* 382.1 (2018), pp. 1–11. DOI: 10.1016/j.physleta.2017.10.027.

G. Iacobello, L. Ridolfi, M. Marro, P. Salizzoni, and S. Scarsoglio. "Complex Network Analysis of Wind Tunnel Experiments on the Passive Scalar Dispersion in a Turbulent Boundary Layer". In: *Progress in Turbulence VIII*. ed. by R. Örlü, A. Talamelli, J. Peinke, and M. Oberlack. Cham: Springer International Publishing, (2019), pp. 215–220. DOI: 10.1007/978-3-030-22196-6_34.

G. Iacobello, M. Marro, L. Ridolfi, P. Salizzoni, and S. Scarsoglio. "Experimental investigation of vertical turbulent transport of a passive scalar in a boundary layer: Statistics and visibility graph analysis". In: *Physical Review Fluids* 4.10 (2019), p. 104501. DOI: 10.1103/PhysRevFluids.4.104501.

5.1 Motivation

Time-series analysis is a broadly adopted approach to study the temporal evolution of dynamical systems, specifically those with high intrinsic complexity. So far, different methods (such as PDFs or Fourier and wavelet transforms) have been developed to extract information from time-series. However, since each method unavoidably loses some information about the temporal structure of the series analysed, new approaches are continuously required to fill this lack.

In this work, the visibility algorithm [97] is employed to systematically investigate time-series from wall-bounded turbulent flows, with the aim to advance the

level of information of classical statistics. Motivated by the increasing availability of numerical and experimental measurements, a powerful tool relying on complex networks is provided to extract non-trivial information from time-series. In this way, a different perspective on (wall-bounded) turbulence as well as benchmark results for future network-based investigations are given.

Although the visibility algorithm has been widely employed, the meaning of the metrics extracted from visibility networks has often been merely interpreted as a network feature rather than a time-series feature. For example, a high or low value of average degree (see Eq. (3.3)) in a visibility network indicates the activation of a large or low number of connections, but no information is given on *how* the signal is made to produce higher or lower value of degree. Therefore, although the overall features of the time-series are inherited by the corresponding visibility graphs, it is not straightforward how topological network metrics are affected by different temporal behaviours of the series. To address this issue, in this work, particular care is given not only to relate the network topology to the temporal structure of the series, but also paying attention to the physical interpretation of the network metrics with respect to the flow dynamics. New insights into how the network topology is affected by important temporal features of the mapped signal – namely, irregularities and extreme events – are thus provided. A systematic approach to highlight temporal features of the time-series through specific combinations of the trend of the network metrics is thus proposed and discussed. By doing so, a qualitative correspondence between the network metrics and the flow dynamics is presented, underlying the ability of the method to identify different flow regions.

In order to assess the potential of visibility networks to shed light on the temporal structure of the signals and, in turn, in the turbulence dynamics, two flow setups are selected. Specifically, a numerically simulated (fully-developed) turbulent channel flow and the dispersion of a passive scalar plume in an experimental turbulent boundary layer are focused. In so doing, both numerically and experimentally measured time-series are considered, thus making the visibility algorithm a suitable to map data from either numerical simulations or experiments.

The Chapter is organized as follows. In Section 5.2, the meaning of visibility-based network metrics is related to the temporal structure of the signals. The application of the natural visibility algorithm to velocity time-series in a turbulent channel flow is discussed in Section 5.3. The behaviour of the network metrics as a function of the wall-normal coordinate and their relation to the flow dynamics is emphasized for the three velocity components. Additionally, in Section 5.3.3, a parametric analysis on the temporal discretization as well as an exploratory study on the application of the horizontal visibility to time-series of streamwise velocity are also provided. Finally, in Section 5.4, the visibility algorithm is exploited to study the dispersion of a passive scalar plume emitted by an elevated source in a turbulent boundary layer. Experimental measurements of velocity and passive

scalar concentration are performed in a turbulent boundary layer over a rough-wall, which is intended to represent the dispersion process of a passive scalar in the atmospheric boundary layer. In particular, classical statistics of concentration and turbulent transport are given in Section 5.4.1, while the behaviour of network metrics are provided in Sections 5.4.2 and 5.4.3.

5.2 Time-series structure in the view of network metrics

Since its introduction for mapping time-series into networks, the visibility algorithm has succeeded in a large variety of research fields, owing to its ability to preserve structural information of time-series in the topology of the associated visibility graphs [121, 174]. For example, periodic time-series are converted into regular networks, namely graphs where nodes have constant degree centrality values related to the periods of the series. Uncorrelated and random series, instead, are mapped into exponential random graphs, that is, random networks with decreasing exponential degree distributions [97, 124].

In general, if two different time-series are compared, they can differ in several ways. In this analysis, we focused on the presence of *peaks* and *irregularities*, mainly for two reasons. First, the occurrence and temporal collocation of extreme events (i.e., peaks) and irregularities represent some of the fundamental features to characterize turbulent flows. Second, the visibility graph is a suitable method to evidence this kind of flow properties and translate them into the network topology [175]. More in detail, a point of a time-series, $s(t_i)$, is said a *peak* if it is a local (or global) maximum of $s(t_i)$, with order of magnitude comparable with the maximum excursion of the series, $\Delta s = (\max_i [s(t_i)] - \min_i [s(t_i)])$. Peaks generally have higher probabilities to connect to other points in the series, because obstacles to the visibility are avoided from higher positions. Therefore, the most important nodes (i.e., the *hubs*) for a visibility-network are associated with positive peaks in the series, because very high values are more likely to see other nodes (i.e., hubs have a better visibility). On the other hand, *irregularities* are temporal variations with order of magnitude much smaller than Δs, and defined as local barriers decreasing the visibility of the surrounding points. To grasp the concept, Figure 5.1 shows four basic examples of time-series and the corresponding networks, $\mathcal{G}$. Time-series $s(t_i)_{(a),(c)}$ (Figure 5.1(a),(c)) are clearly more regular than the series $s(t_i)_{(b),(d)}$ (Figure 5.1(b),(d)). Moreover, $s(t_i)_{(c)}$ and $s(t_i)_{(d)}$ have three peaks (for $t_i = 1, 50, 100$) while series $s(t_i)_{(a)}$ and $s(t_i)_{(b)}$ display two peaks (for $t_i = 1, 100$). Therefore, while the networks $\mathcal{G}_{(a)}$ and $\mathcal{G}_{(c)}$ are well organized in clusters (one cluster for $\mathcal{G}_{(a)}$, two for $\mathcal{G}_{(c)}$), $\mathcal{G}_{(b)}$ and $\mathcal{G}_{(d)}$ appear more complex. This happens because in $\mathcal{G}_{(b),(d)}$ there are many nodes with low visibility due to the presence of irregularities.

It must be emphasized that we refer to peaks as the local (or global) highest

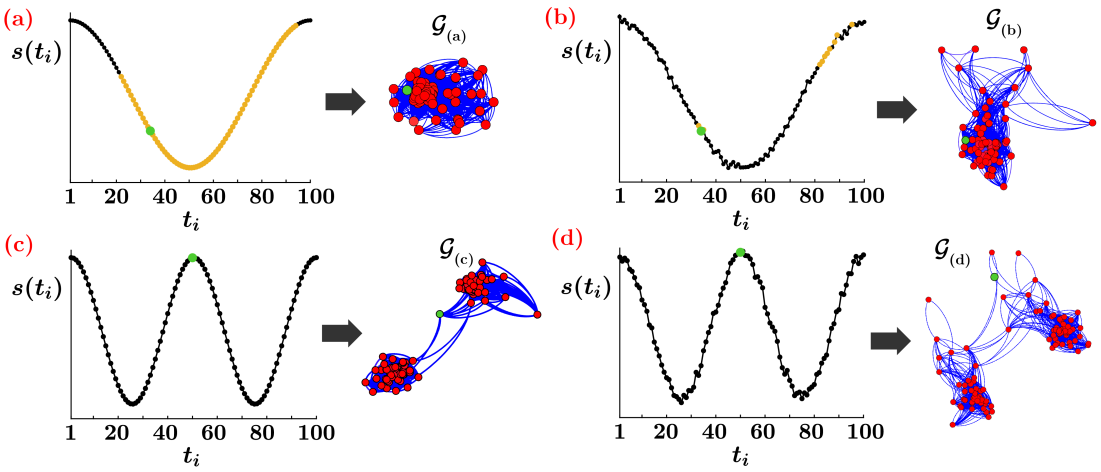


Figure 5.1: Four examples of time-series with different temporal features and corresponding visibility network representations (obtained through the *OpenOrd* layout algorithm [130]). In panels (a) and (b), the green-coloured dot indicates the point $s(t_i = 34)$, while yellow points highlight its first neighbourhood, Γ^1_{34}. In panels (c) and (d), the green-coloured dot evidences the point $s(t_i = 50)$.

positive values in the series, but this does not necessarily imply that peaks also correspond to *outliers*, i.e. very large values with respect to a local subset of data. On the contrary, outliers typically correspond to peaks. To clarify this concept, in Figure 5.2, two simple examples of time-series (illustrated as black stems) are mapped into visibility networks, where nodes and links are depicted as red dots and blue lines, respectively. In the illustrative time-series of Figure 5.2(a), outliers are not present and peaks correspond to nodes $t_i = \{2, 4, 7, 9\}$; in Figure 5.2(b), instead, peaks correspond to outliers, namely nodes $t_i = \{2{,}15\}$. Furthermore, it is worth highlighting that the visibility criterion emphasizes the positive peaks, but not the negative ones. Consequently, when the series mainly display *pits* (i.e., negative peaks) instead of positive peaks, it is possible to exploit the Eq. (4.1) (or the Eq. (4.2) for the horizontal visibility) to build visibility networks from the complementary series, $-s(t_i)$. The comparison of the metrics extracted from the original series, $s(t_i)$ and its opposite, $-s(t_i)$, allows one to characterize the peak-pit asymmetry in the series [176], namely if peaks are mainly positive or negative.

Among all the metrics available to analyse complex networks, the (average) mean link-length, $\overline{d^t}$, the assortativity coefficient, r, and the transitivity, $\mathcal{T}$, turned out to be the metrics that better capture the temporal structure of the time-series in terms of peaks and irregularities, inheriting important features of the turbulent flow dynamics. In what follows, details on each metric and its relation to the temporal structure of the signals are provided, by highlighting the metric meaning in terms of peaks and irregularities.

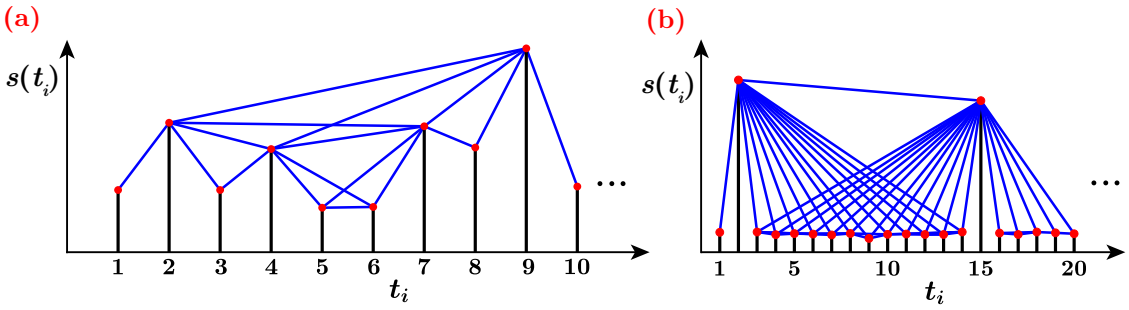


Figure 5.2: Examples of two intervals of time-series, (t_i, s_i), and corresponding visibility networks highlighting (a) peaks and (b) outliers. Nodes and links are depicted as red dots and blue lines, respectively. (a) First 10 observations of a series extracted from a uniform probability distribution in the interval [0,1]. (b) First 20 observations of a series extracted from a uniform probability distribution in the interval [0,1], with periodic spikes (every 14 instants, t_i) uniformly distributed in the interval [0,100]. The values of s_i in the vertical axis are not shown due to the invariance of visibility algorithm to affine transformations.

5.2.1 Transitivity analysis

The transitivity, $\mathcal{T}$, is the first metric focused (see Eq. (3.8)). We recall that each pair of nodes $(j,k) \in \Gamma^1(i)$ (where $\Gamma^1(i)$ is the first neighbourhood of node i) always forms a connected triple with node i (e.g., see Figure 3.6). Therefore, the total number of triples in the network depends on the size of all the neighbourhoods $\Gamma^1(i)$, because more neighbours a node has, more triples it activates. On the other hand, triangles are formed only if the nodes $(j,k) \in \Gamma^1(i)$ are also linked to each other, that is, if $\mathcal{A}_{i,j} = \mathcal{A}_{i,k} = \mathcal{A}_{j,k} = 1$. In general, connections between nodes separated by short time intervals are the most probable ones, because (differently from random signals) time-series of physical quantities are not expected to sharply change in time. Therefore, nodes that are close in time are more likely to form triangles. If two neighbours $(j,k) \in \Gamma^1(i)$ are far in time, instead, there are many nodes in between j and k so that there is a high probability to find a node that obstructs the inter-visibility of j and k. As a result, the total number of triangles and triples, and therefore the transitivity, strongly depend on the inter-visibility of nodes inside each neighbourhood.

The transitivity can be then interpreted as a measure to characterize the typical convexity properties of a signal on some small-to-intermediate time-scales [177]. For example, in Figure 5.1, any point in the ranges $t_i = (1-50)$ and $t_i = (51-100)$ of $s(t_i)_{(c)}$ (see Figure 5.1(c)) has basically the same inter-visibility of corresponding points in $s(t_i)_{(a)}$ (see Figure 5.1(a)), namely there are no substantial local changes of regularity between $s(t_i)_{(a)}$ and $s(t_i)_{(c)}$. As a consequence, the value of transitivity of $\mathcal{G}_{(a)}$ and $\mathcal{G}_{(c)}$ are expected to be scarcely affected by different occurrence of peaks.

To show that, Figure 5.3(a) reports the values of transitivity (green bars) for the four visibility network shown in Figure 5.1. As evident, $\mathcal{T}_{(a)}$ and $\mathcal{T}_{(c)}$ are actually almost the same, being equal to 0.72 and 0.73, respectively. The presence of more (or less) peaks in a time-series then does not significantly modify the short-term inter-visibility (i.e., the transitivity) of nodes. In fact, also $\mathcal{T}$ values of $\mathcal{G}_{(b)}$ and $\mathcal{G}_{(d)}$ are almost equal (see the $\mathcal{T}$ values in Figure 5.3(a)), being the irregularities of time-series $s(t_i)_{(b)}$ and $s(t_i)_{(d)}$ very similar.

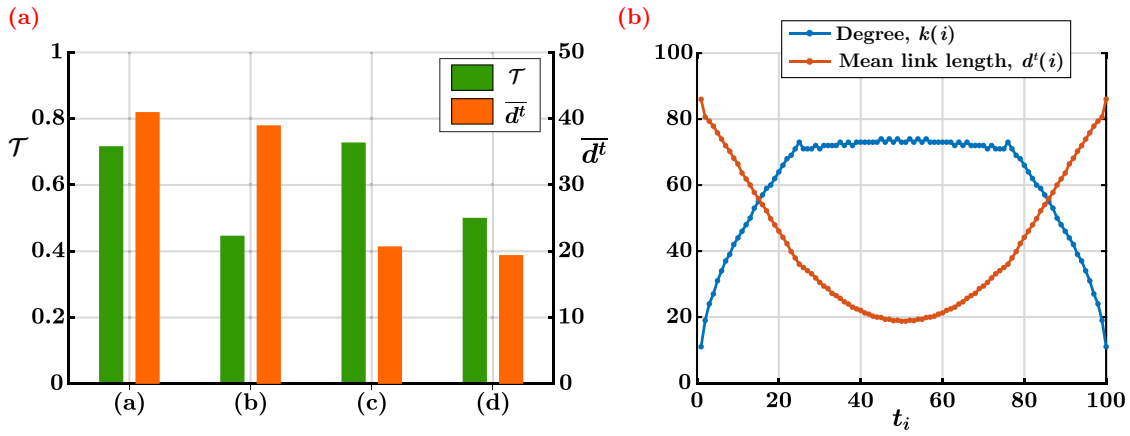


Figure 5.3: (a) Bar plot of the transitivity (green) and average mean link-length (orange) values for the four time-series of Figure 5.1. (b) Degree centrality and mean link-length values of the nodes of the network shown in Figure 5.1(a).

On the contrary, the time-series $s(t_i)_{(b),(d)}$ clearly display irregularities if compared with time-series $s(t_i)_{(a),(c)}$. The inter-visibility among neighbours of a generic node is obstructed because of the irregularities in the time-series. As a paradigmatic example, we consider an arbitrary node, e.g. $i = 34$, and its first neighbourhood, $\Gamma^1(34)$, highlighted as a green-coloured and orange dots, respectively, in Figure 5.1(a),(b). While in $\mathcal{G}_{(a)}$ the neighbourhood $\Gamma^1(34)$ includes either short-term, medium-term, and long-term links, in $\mathcal{G}_{(b)}$ the neighbourhood $\Gamma^1(34)$ includes only short-term and long-term connections. Therefore, the number of triangles (relative to the number of triples) in which is involved a generic node (e.g., $i = 34$) is generally lower in irregular networks than in regular ones. As a result, the values of transitivity $\mathcal{T}_{(b)}$ and $\mathcal{T}_{(d)}$ are much lower than $\mathcal{T}_{(a)}$ and $\mathcal{T}_{(c)}$, as observed in Figure 5.3(a).

Summarizing, the transitivity is much more affected by local variations due to the presence of irregularities rather than the presence of local peaks in the series. In terms of flow dynamics, the transitivity is related to the presence of local fluctuations between consecutive peaks. Recalling that the visibility algorithm is insensitive to the mean and standard deviation values of the mapped signals, the transitivity is thus a net measure of the intrinsic fluctuation level of the time-series.

5.2.2 Mean link-length analysis

The second metric considered is the mean link-length, $\overline{d^t}$, (where the $\overline{\bullet}$ notation indicates the average over all network nodes, see Eq. (3.13)) or equivalently its inverse metric that is the average peak occurrence, $\phi = 1/\overline{d^t}$. If peaks often occur in a series (as in Figure 5.1(c),(d)), points far from each other are not visible because far connections are hampered by peaks, and $\overline{d^t}$ is consequently strongly reduced. The visibility of a generic node in the networks $\mathcal{G}_{(c)}$ and $\mathcal{G}_{(d)}$ is limited by the peak at $i = 50$ (green-coloured dot in Figure 5.1(c),(d)), which in turn divides the networks into two main clusters (see the network representations in Figure 5.1(c),(d)). The value of $\overline{d^t}_{(c)}$ and $\overline{d^t}_{(d)}$ are indeed much lower than $\overline{d^t}_{(a)}$ and $\overline{d^t}_{(b)}$, respectively, as reported in Figure 5.3(a). On the other hand, $\overline{d^t}$ is not essentially affected by the irregularities of a series. Irregularities mostly prevent medium-term connections than short and long-term links but, averaging over all nodes in the network, a value of the order of medium-term links is generally obtained for $\overline{d^t}$. In fact, in Figure 5.3(a) the value of $\overline{d^t}_{(b)}$ is approximately equal to $\overline{d^t}_{(a)}$, and $\overline{d^t}_{(d)}$ is almost the same of $\overline{d^t}_{(c)}$, indicating that there are no relevant changes in the global mean link-length due to irregularities. Therefore, large d^t values correspond to hubs in the network and peaks in the series, as peaks activate long-range links (i.e. farther temporal *horizons*).

It should be noted that, although the degree centrality, $\mathcal{K}$, is usually adopted as the metric to characterize hubs, the mean link-length reveals to be a more reliable metric than degree for the temporal characterization of peaks in visibility networks. For example, in the series shown in Figure 5.1(a) (which is a bowl-like series with a minimum value), the nodes in the range $t_i = (25-75)$ have maximum degree values as illustrated in Figure 5.3(b), but they are not the peaks of the series (which are nodes $t_i = 1, 100$). This is due to the fact that the series reported in Figure 5.1(a) is fully convex only in the range $t_i = (25-75)$, while the long-range visibility of peaks (i.e., $t_i = 1, 100$) is obstructed by the points in the surroundings of peaks (e.g., nodes at $t_i = (2-25)$ or $t_i = (75-99)$), thus creating local barriers to the visibility of lower points of the series. On the contrary, as illustrated in Figure 5.3(b), the mean link-length is able to capture the presence of peaks in the series, since the highest $d^t(i)$ values correspond to the maximum values in the series. In visibility networks, therefore, hubs can be better identified as nodes with long-range horizons (i.e., high d^t values), rather than nodes with high degree centrality. It is also worth noting that while the kurtosis is an estimation index of extreme values in a PDF, the mean link-length quantifies the average temporal distance between extreme events (while the PDF is invariant to the temporal structure of the signal).

To conclude, the (average) mean link-length, $\overline{d^t}$, (or $\phi = 1/\overline{d^t}$) is strongly influenced by the occurrence of peaks in time (horizontal separation), being slightly affected by the irregularities. The mean link-length represents a characteristic temporal distance between two visible data in a series (while ϕ is a characteristic

frequency), thus indicating how isolate and sporadic extreme events are. In particular, low $\overline{d^t}$ (i.e., high ϕ) values are obtained when the recurrence of peaks is high.

5.2.3 Assortativity analysis

Although the mean link-length (or the average peak occurrence) is sensitive to the appearance of peaks in time (horizontal separation), it does slightly take into account the relative intensity of peaks compared to all the other values in the series (vertical separation).

To address this issue, we also investigated the assortativity coefficient, r (see Eq. (3.9)). We recall that positive r values are obtained when nodes are linked with other nodes of similar degree and the network is said to be assortative, while the network is said to be disassortative if $r < 0$, or non-assortative if $r = 0$. When peaks are focused, the assortativity coefficient quantifies the extent to which peaks (that are expected to have more visibility) are more prominent with respect to small fluctuations (which are expected to have less visibility). Highly positive values of r indicate that peaks are slightly pronounced with respect to the other values in the series (e.g., Figure 5.2(a)), while strongly negative values of r indicate a substantial presence of outliers (e.g., Figure 5.2(b)). Being the degree centrality, $\mathcal{K}$, a direct measure of the visibility of nodes, it can concurrently account for both the recurrence of peaks and the presence of irregularities. In other words, the average degree, $\overline{\mathcal{K}}$, combines the features of both the transitivity, $\mathcal{T}$, and the mean link-length, $\overline{d^t}$, in a single global metric. Therefore, due to its intrinsic definition, the degree variation in general cannot be univocally related to a specific temporal feature (either peaks or irregularities occurrence). However, since $\mathcal{K}$ is able to retain information on both large and small fluctuations, it represents a suitable metric to evaluate the assortativity coefficient (which quantifies the relative intensity between large and small fluctuations).

In conclusion, the assortativity coefficient, r, is a measure of the vertical separations in the series, that is how intense peaks are with respect to the other data in the time-series.

5.2.4 Combining the network metrics

From the metric interpretation provided so far it emerges that the visibility algorithm turns out to be able to capture three main features of the temporal structure of a series: the recurrence and the relative intensity of peaks, as well as the presence of irregularities. Specifically, for the peaks characterization, $\overline{d^t}$ (or ϕ) and r are able to highlight the horizontal and vertical temporal structure of the signals, respectively. By combining the behaviours of the selected (global) metrics,

Table 5.1: Scheme of the ways two time-series, $s_{(1)}$ and $s_{(2)}$, can differ and corresponding behaviours of the global network-metrics, $\mathcal{T}$, $\overline{d^t}$ and r.

Cases	**Temporal structure features**	**Metric behaviours**	**Series $s_{(1)}$**	**Series $s_{(2)}$**
Case A	Peaks occur more frequently in $s_{(2)}$ than in $s_{(1)}$	$\mathcal{T}_{(2)} \approx \mathcal{T}_{(1)}$, $\overline{d^t}_{(2)} < \overline{d^t}_{(1)}$ $(\phi_{(2)} > \phi_{(1)})$		
Case B	The series $s_{(2)}$ is more irregular than the series $s_{(1)}$	$\mathcal{T}_{(2)} < \mathcal{T}_{(1)}$, $\overline{d^t}_{(2)} \approx \overline{d^t}_{(1)}$ $(\phi_{(2)} \approx \phi_{(1)})$		
Case C	Peaks occur more frequently in $s_{(2)}$ than in $s_{(1)}$, and $s_{(2)}$ is more irregular than $s_{(1)}$	$\mathcal{T}_{(2)} < \mathcal{T}_{(1)}$, $\overline{d^t}_{(2)} < \overline{d^t}_{(1)}$ $(\phi_{(2)} > \phi_{(1)})$		
Case D	Peaks occur less frequently in $s_{(2)}$ than in $s_{(1)}$, and $s_{(2)}$ is more irregular than $s_{(1)}$	$\mathcal{T}_{(2)} < \mathcal{T}_{(1)}$, $\overline{d^t}_{(2)} > \overline{d^t}_{(1)}$ $(\phi_{(2)} < \phi_{(1)})$		
Case E	The series $s_{(1)}$ display (positive) peaks while the series $s_{(2)}$ display outliers	$r_{(2)} < r_{(1)}$, $r_{(1)} > 0, \quad r_{(2)} < 0$		

it is possible to provide a comparative temporal characterization of different time-series.

First, the transitivity and the mean link-length are focused. From the perspective of these two metrics, a series can differ from another series through a combination of the metrics behaviours, namely $\mathcal{T}$ and $\overline{d^t}$ can increase, decrease, or remain almost constant. Excluding the last combination, i.e. the one in which both $\mathcal{T}$ and $\overline{d^t}$ are almost constant (i.e., the two compared time-series share the same temporal features), four different *cases* can occur. These four cases are labelled as *A, B, C, D* in Table 5.1, which also reports the metric behaviours and illustrative examples of time-series. Therefore, given the behaviours of $\mathcal{T}$ and $\overline{d^t}$, it is possible to infer from Table 5.1 how time-series differ in terms of irregularities and peak occurrence, respectively.

For each of the four cases *A*-*D* reported in Table 5.1, the assortativity coefficient can be evaluated in order to characterize the vertical separation of the signals. As a rule of thumb, $r = 0$ discriminates between the prominence of peaks (i.e, $r > 0$) and the prominence of outliers (i.e., $r < 0$); this feature is indicated as the case *E* in Table 5.1. If the network is non-assortative (i.e., $r \approx 0$), in general neither peaks nor outliers are expected to be prominent. However, if two signals are compared,

it can be inferred that the signal showing $r \approx 0$ is more likely to have outliers or peaks than the signal with $r > 0$ or $r < 0$, respectively. In terms of extreme events, hence, ϕ or $\overline{d^t}$ are more reliable to detect the occurrence of peaks in the series, while r is able to discern between peaks (in which r is generally positive) and outliers (in which r is generally negative) in a time-series. For instance, in Figure 5.2(a), $\phi = 0.230$ and $r = 0.22$, while in Figure 5.2(b) $\phi = 0.156$ and $r = -0.19$. These values are evaluated as described in the caption of Figure 5.2. Moreover, large values of ϕ indicate that the corresponding series has many peaks, which appear as outliers only if r decreases towards negative values.

To summarize, the metrics, $\mathcal{T}$, ϕ (or $\overline{d^t}$) and r, should be analysed in pairs in order to infer the temporal structure of time-series evaluated at different spatial locations.

5.3 Analysis of velocity time-series in turbulent channel flows

In this Section, the metric interpretation about the temporal structure of the signals provided so far is exploited to investigate velocity time-series in a turbulent channel flow at $Re_\tau = 1000$ (full details of the DNS are reported in Appendix B.2). Velocity time-series are adopted to build the corresponding networks as the velocity field is one of the most intuitive quantity to characterize a fluid flow. However, the visibility graph method can be applied to other quantities of turbulence interest, such as the Reynolds shear stress, the kinetic energy, or the vorticity field. In order to illustrate the potential of visibility networks to inherit the temporal structure of a signal, the streamwise velocity component, u, is firstly focused in Section 5.3.1. A comparison of the results for the three velocity components, (u, v, w), is given in Section 5.3.2, while complementary results are provided in Section 5.3.3.

It should be emphasized that, in what follows, the metric behaviours are only shown as a function of the wall-normal coordinate, y^+, as the metrics are averaged over the homogeneous directions, x and z. In fact, y is the only direction where spatial inhomogeneities develop, while the velocity field is statistically homogeneous along x and z. Since network measures inherit the properties of the mapped time-series, also the global metrics (i.e., averaged over the nodes of each network) are assumed to be statistically homogeneous in the x-z directions. Therefore, the results provided in this section (in terms of $\overline{d^t}$, $\mathcal{T}$, r and $\overline{\mathcal{K}}$) are intended to be averaged in the homogeneous directions without any further specification.

5.3.1 Wall-normal behaviour of network metrics

The behaviour of the network metrics as a function of y^+ for the streamwise velocity time-series is shown in Figure 5.4. Each metric substantial varies along

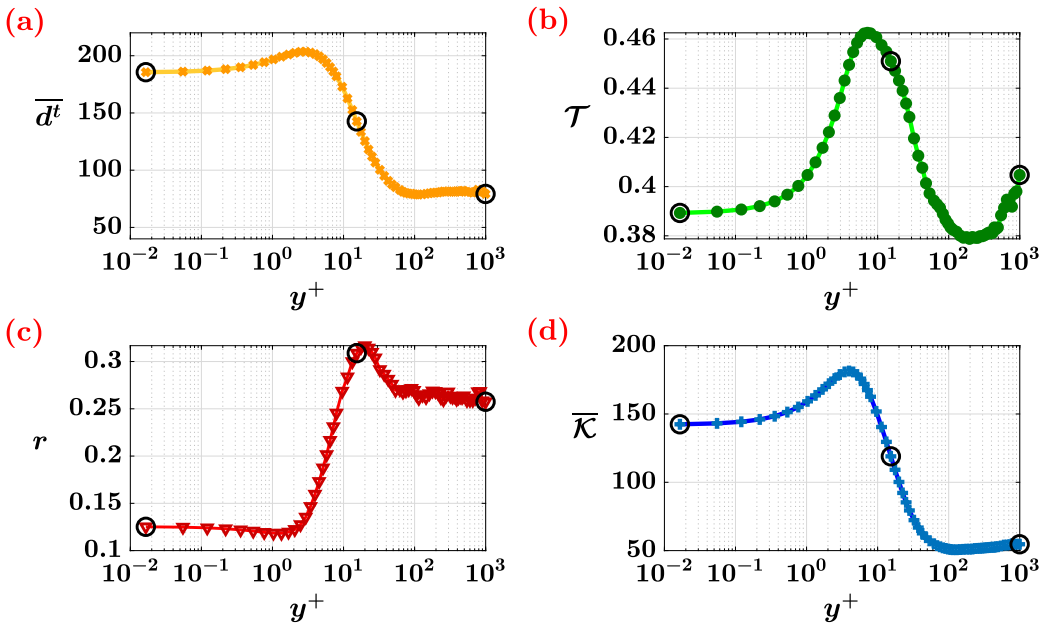


Figure 5.4: Wall-normal behaviour of the average metrics for u in the turbulent channel flow, reported in a log-linear plot. (a) Mean link-length; (b) transitivity; (c) assortativity coefficient; (d) degree centrality. Three representative values at $y^+ = \{0.017, 15.4, 996.3\}$ are also highlighted as black circles.

the wall-normal direction, exhibiting clear and regular trends. The four metrics have overall similar behaviours, rising from the wall up to a maximum value, then decreasing and, finally, barely changing towards the center of the channel. In more detail, the maximum values of the metrics are not exactly at the same y^+ values, but they are quite close in the range $y^+ \approx 4 - 20$. To infer the temporal structure of time-series along the wall-normal coordinate, we start from time-series close to the wall and then proceed towards the center of the channel. In particular, we focus on three representative y^+ stations, i.e. $y^+ = \{0.017, 15.4, 996.3\}$ (highlighted with black circles in Figure 5.4). To better illustrate the network meaning, three time-series are also extracted at the selected y^+ stations; the series are shown in Figure 5.5 with their corresponding graphical (topological) representation. In particular, in order to compare the signals at different y^+, the time-series of the streamwise velocity, u, are normalized by subtracting their mean value and dividing by their standard deviation (this normalization is indicated as u^*).

Moving from $y^+ = 0.017$ to $y^+ = 15.4$, Figure 5.4 shows that the transitivity, $\mathcal{T}$, and the assortativity coefficient, r consistently increase, while the average mean link-length, $\overline{d^t}$, and the average degree, $\overline{\mathcal{K}}$, decrease. This combination of metrics corresponds to the cases *D-E* in Table 5.1 (where $s_{(2)}$ and $s_{(1)}$ correspond to time-series at $y^+ = 0.017$ and $y^+ = 15.4$, respectively). A time-series extracted at

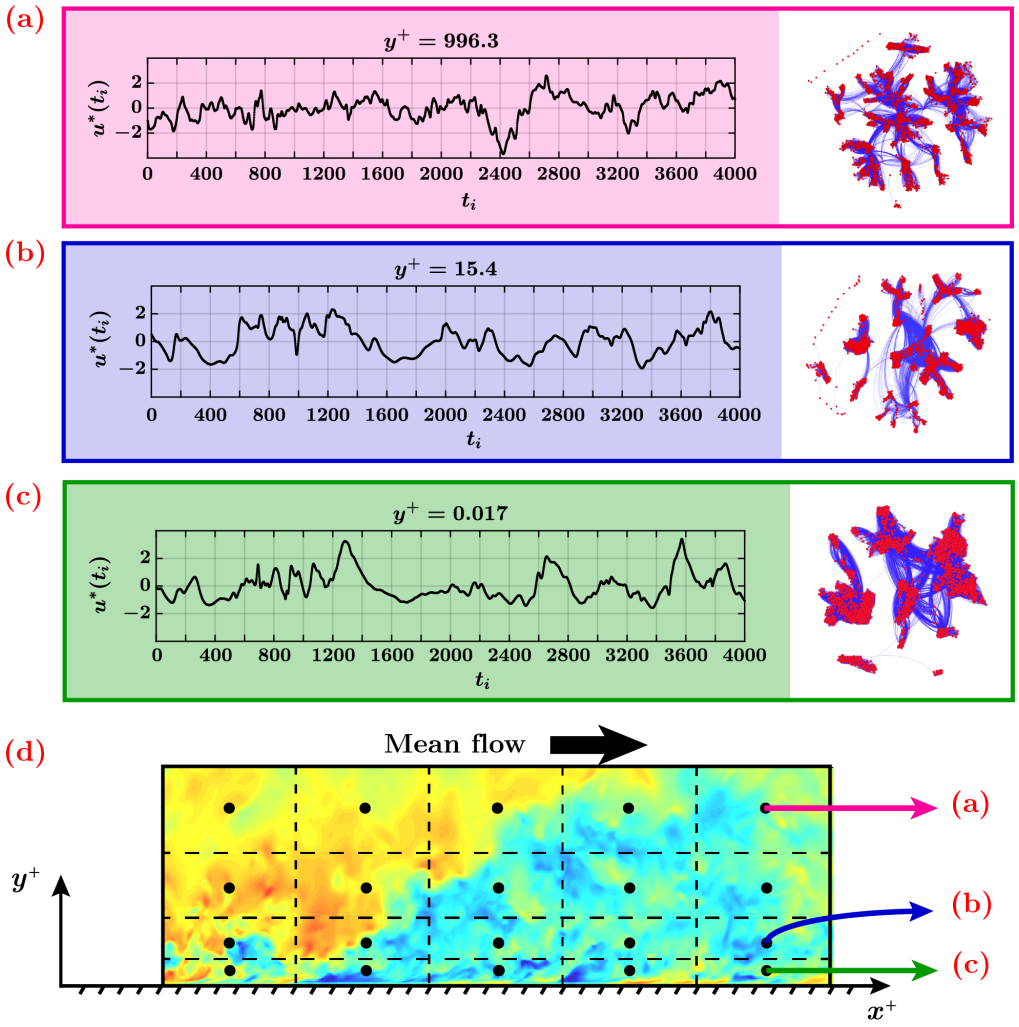


Figure 5.5: (a)-(c) Normalized time-series, u^*, and corresponding graphical representation at the grid-points $X_i = 1601$, $Z_i = 750$ and $y^+ = \{0.017, 15.4, 996.3\}$. The choice of the coordinate in the homogeneous directions, x and z, is arbitrary. (d) Sketch of the channel setup in a (x-y) plane; colours refer to values of streamwise velocity fluctuations, u'.

$y^+ = 15.4$ is then expected to be (on average) more regular than a series extracted at $y^+ = 0.017$ (indicated by the growth of $\mathcal{T}$), with a more frequent occurrence of peaks (indicated by the drop of $\overline{d^t}$), which appear more as peaks than outliers (indicated by the increase of r). The reduction of the average degree, $\overline{\mathcal{K}}$, suggests that the increasing occurrence of peaks affects the global visibility more than the reduction in the irregularities. Looking at the time-series extracted at $y^+ = 15.4$ (Figure 5.5(b)), it is indeed with more peaks than the series extracted at $y^+ = 0.017$

(Figure 5.5(c)). This aspect is also evident in a more clustered topology of the network built on the time-series at $y^+ = 15.4$ (see network representations in Figure 5.5(b)-(c)). Although the regularities appear globally similar, a zoomed view of the signals reveals that the time-series at $y^+ = 0.017$ appears more irregular, as indicated by the transitivity.
Concerning the vertical separation, the assortativity coefficient is always positive (see Figure 5.4(c)), thus indicating that outliers are not likely to appear (as evidenced in the time-series of Figure 5.5(a)-(c)). Moreover, the maximum of assortativity coefficient, r, is found at $y^+ \approx 20$, which is at an intermediate location between the minimum kurtosis value (that is at $y^+ \approx 15$) and the local minimum of skewness (that is at $y^+ \approx 30$) [178]. The assortativity coefficient as well as the mean link-length, therefore, are able to retain combined information on higher-order statistics.

From $y^+ = 15.4$ to $y^+ = 996.3$ (i.e., close to the center of the channel, H), all the average metrics decrease. The specific combination of transitivity and mean link-length corresponds to the case C in Table 5.1 (as both metric values diminish). Accordingly, we expect that a time-series extracted at the center of the channel is (on average) less regular than a series at $y^+ = 15.4$ and with a more frequent recurrence of peaks. This behaviour can be clearly seen in Figure 5.5 where the time-series at the center of the channel (Figure 5.5(a)) is more fluctuating than the time-series at $y^+ = 15.4$ (Figure 5.5(b)), and the corresponding network appears more clustered and disordered. It is interesting to note that from $y^+ \approx 10^2$ to the center of the channel, all the four metrics barely change. Therefore, most of the variations of the network structure are present in the inner layer ($y^+ \leq 10^2$), namely the most affected region by the presence of the wall.

It is important to remark that the behaviour of a single metric is not a sufficient information, but a combination of two metrics, e.g. $(\mathcal{T}, \overline{d^t})$ or $(r, \overline{d^t})$, instead, determines how two time-series differ in terms of irregularities and recurrence of peaks or vertical separation and horizontal separation, respectively. Moreover, we do not refer to the specific value assumed by the metric, but the analysis is comparative as it focuses on the trend each metric assumes as a function of the distance from the wall. Specifically, comparing a time-series at the wall to one at the center of the channel, peaks are expected to be remarkably closer for $y = H$, while irregularities do not substantially change ($\overline{d^t}$ decreases while $\mathcal{T}$ slightly increases). In fact, as shown in Figure 5.5(a), in the center of the channel peaks occur more frequently but the irregularity between them remains basically unvaried. However, this trend is not monotonic along y^+, since the time-series locally (around $y^+ = 15.4$) change their regularity. In terms of the network topology, close to the wall the network is composed by different sub-networks (see Figure 5.5(c)), corresponding to the peaks of the series, which are widely connected with each other and internally. Going towards $y^+ = 15.4$ (see Figure 5.5(b)), the simultaneous decrease of $\overline{d^t}$ and increase of $\mathcal{T}$ mainly break down long connections among the sub-networks. The

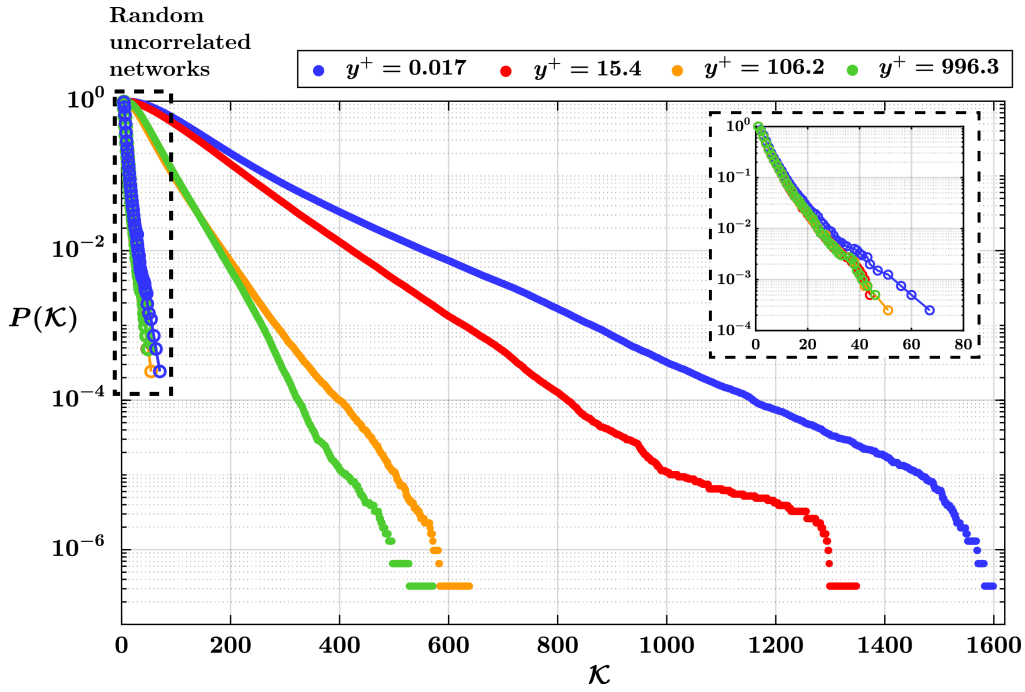


Figure 5.6: Cumulative degree distributions, $P(\mathcal{K})$, for the u-based visibility network, averaged over the homogeneous directions. The distributions close to the vertical axis (highlighted in the inset) correspond to networks built on shuffled time-series (similar colours refers to similar y^+ values). The slopes of the distributions are $(-1.01, -1.08, -2.85, -2.86) \cdot 10^{-2}$ from $y^+ = 0.017$ to $y^+ = 996.3$, respectively; the slope for the shuffled series is about $2 \cdot 10^{-1}$. The coefficient of determination, R^2, of the fittings is always above 0.99.

drop of $\overline{d^t}$ plays a major role here, acting to split long-term links. The subsequent decrease of both $\overline{d^t}$ and $\mathcal{T}$ (from $y^+ = 15.4$ to 996.3) breaks principally intra-network connections. At this stage, the prevailing effect is locally induced by the increase of irregularity, which leads to a ramification of each sub-network (see Figure 5.5(a)).

Degree centrality and its probability distribution. Although it is just used as a measure of the relative intensity of irregularities and peaks, a comment on the degree centrality can be eventually carried out. Since the average degree is proportional to the link density in an undirected network (see Eq. (3.4)), a high value of $\overline{\mathcal{K}}$ indicates a globally convex time-series, while low values indicate a strong fragmentation of the visibility network [177]. As a result, considering the behaviour of $\overline{\mathcal{K}}$ in Figure 5.4(d), towards the channel center the time-series are globally more fragmented than the time-series close to the wall, confirming what found observing the trends of $\mathcal{T}$ and $\overline{d^t}$. Besides, the degree centrality can be effectively exploited to qualitatively estimate the (linear and non-linear) temporal dependences existing

in a time-series, by means of the degree probability distribution [122]. Figure 5.6 shows the cumulative degree distributions, $P(\mathcal{K})$, in a semi-log plot for networks built at four representative y^+. As for the metrics shown in Figure 5.4, the degree distribution is computed for each network at each grid location, and all the computed distributions are then averaged over the homogeneous directions. As evident from the Figure 5.6, the tail of the distributions reveals decreasing exponential trends, namely the highest degree values (corresponding to the network hubs) are generally very infrequent. In particular, the exponent of the fitting of the $P(\mathcal{K})$ increases (in modulus) from the wall towards the center of the channel, $y^+ = 1000$. This is consistent with the temporal integral scale measurements [179], which decrease from the wall to the center of the channel. However, it should be noted that, differently from the horizontal visibility algorithm, the behaviour of the degree distribution for the natural visibility algorithm also depends on the PDF of the mapped time-series [123]. Therefore, in order to isolate (from the PDF contribution) the net impact of (linear and non-linear) dependences in the turbulent time-series, we built four series by shuffling four velocity time-series (at arbitrary (x, z) locations) at the same wall-normal distances considered, i.e., $y^+ = \{0.017, 15.4, 106.2, 996.3\}$. As shown in Figure 5.6 (and highlighted in the inset), the slopes of $P(\mathcal{K})$ from the shuffled series are substantially steeper than the turbulent time-series. This demonstrates the key role of the (linear and non-linear) correlation aspects of the turbulent series.

Summary and network metrics in the view of the flow dynamics. In summary, we have pointed out the ability of the visibility-based networks to shed light on the temporal structure of the corresponding mapped time-series. Indeed, the temporal features of the series are actually as predicted by combining the network metrics. Through the behaviour of the metrics along the y^+ direction, Figure 5.4 yields first important results on the presence, dislocation and intensity of extreme events and irregularities of the time-series. This kind of information can enrich the comprehension of the flow dynamics. To this purpose, now we try to relate the network metrics with the turbulent flow dynamics, which is responsible for the time-series behaviour. Looking at the Figure 5.4, three regions are particularly interesting (i.e. $y^+ \lesssim 7$, $7 \lesssim y^+ \lesssim 150$, and $y^+ \gtrsim 150$) where the average metrics mostly change their trend. It should be noted that the values of y^+ delimiting such regions are very close to the limit values, $y^+ = 5$ and $y^+ = 100$, of the viscous sub-layer and inner layer, respectively. The region for $y^+ < 5$ is characterized by slow moving fluid and the flow dynamics are dominated by the viscous shear stresses. The normalized time-series $u^*(t_i)$ here can be assumed to roughly share a similar temporal structure (although their mean and standard deviation values clearly change along y^+). The corresponding metrics (see Figure 5.4) highlight this behaviour resulting in barely increasing trends. As previously observed, around $y^+ \simeq 4 - 7$ (which is the upper bound of the viscous sub-layer) the transitivity,

mean link-length and degree reach their maximum values. In the range $y^+ \simeq 4-7$ we expect, in terms of time-series shape, a minimum number of peaks along with the minimum irregularities. Recalling that all signals are normalized with the local mean and standard deviation, a possible interpretation is the following. Around $y^+ \simeq 4-7$, we are approaching the buffer layer ($5 < y^+ < 30$), i.e., an intermediate region where viscous shear stress starts decreasing while turbulence activity begins to grow. However, at the very beginning ($y^+ \simeq 4-7$), turbulent processes are very low, thus resulting in a minimum of irregularities that act over a signal that is still affected by slow temporal variations (i.e., low number of peaks). The combination of these dynamics reasonably explains the maxima reached by the three metrics ($\mathcal{T}$, $\overline{d^t}$ and $\overline{\mathcal{K}}$) around the region $y^+ \simeq 4-7$. The maximum of the assortativity coefficient, instead, is located within the buffer layer, where the turbulence activity is strong and the vertical separation in the signal is low (recall that the kurtosis of u is minimum within the buffer layer). For $y^+ \gg 5$ the flow dynamics are more affected by the Reynolds stresses, and the flow shows a tendency to organize into coherent turbulent patterns. The structure of the time-series is then affected by turbulent processes (such as ejections and sweeps), leading to rapid temporal variations. This behaviour could be recognized in the drop of the average metrics in Figure 5.4. As y^+ further increases ($y^+ > 100$), the turbulent patterns are less affected by the wall and they can develop in larger structures. However, the coexistence of multiple scales and the more complicated flow structure seems not to translate into a clear trend for the network metrics.

5.3.2 Comparison of network metrics for different velocity components

In order to extend the analysis of the visibility networks to all the velocity components, (u, v, w), the wall-normal behaviour of the four selected metrics is shown in Figure 5.7.

In general, the mean link-length, $\overline{d^t}$, the assortativity coefficient, r, and average degree, $\overline{\mathcal{K}}$, measured on the time-series of v and w show trends similar to those of u, while different trends are obtained considering the transitivity, $\mathcal{T}$. Additionally, the maximum values of $\overline{d^t}$, r and $\overline{\mathcal{K}}$ shift towards higher y^+ coordinates if the velocity components u, w and v are focused, respectively.
In more detail, the trends over y^+ of $\overline{d^t}$ (see Figure 5.7(a)) for the three velocity components are similar, but values for the streamwise velocity, u, are overall higher than those displayed by v and w. The relative difference decreases towards the center of the channel. The scenario for the assortativity coefficient, r (see Figure 5.7(c)), and the degree centrality, $\overline{\mathcal{K}}$ (see Figure 5.7(d)), is analogous to the $\overline{d^t}$. Differences for the $\overline{d^t}$, r and $\overline{\mathcal{K}}$ values of the three components are marked close to the wall, while the metric values tend to coincide approaching the channel center. This behaviour can be explained by considering that close to the wall the presence

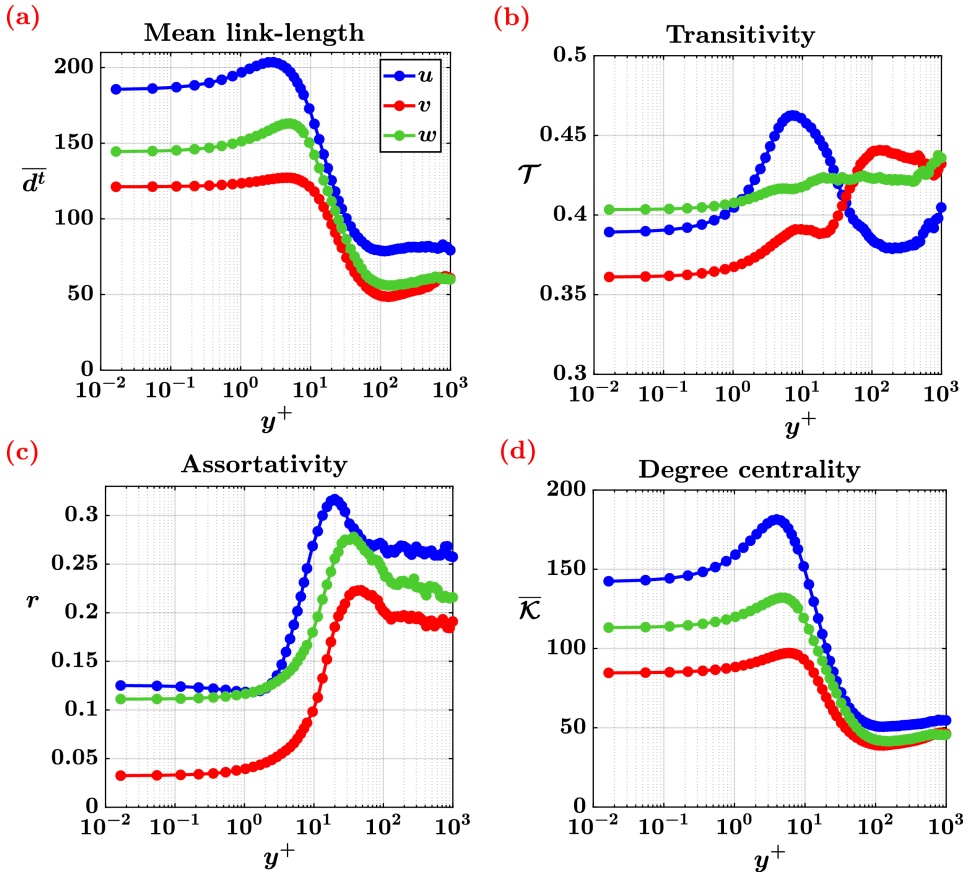


Figure 5.7: Average profiles of (a) mean link-length, (b) transitivity, (c) assortativity coefficient, and (d) degree centrality, evaluated from time-series extracted from the velocity field, (u, v, w).

of the wall itself strongly influences and differently characterizes the flow dynamics in the three directions of the velocity, and consequently the networks based on the corresponding time-series are affected. It is remarkable to observe that the values of assortativity for the networks built on the v component are much lower close to the wall than the corresponding values of u and w. This is in line with the great difference in the kurtosis levels between different velocity components, which is stronger for v and less intense for w and u (e.g., see Ref.s [178, 180]). On the contrary, the wall effects decrease moving far from the wall ($y^+ > 100$), thus differences among the metrics built on u, v and w, reduce.

Focusing on the transitivity, $\mathcal{T}$, the metric difference among velocity components is even more accentuated. In fact, in the region $y^+ < 100$, not only values are different but also metrics display different trends. In particular, the wall-normal velocity component, v, is strongly affected by the presence of the wall (recall that close to the wall the motion corresponds to flow in planes parallel to the wall [4]) and

this in turn involves the transitivity. For example, spikes with large negative values can be found in the time-series of v as a consequence of strong events that appear only in the very near-wall region, revealed by high kurtosis levels [180]. Since these deep peaks are negative and relatively short, the degree and the mean link-length of the corresponding networks are barely affected, while the transitivity and the assortativity are strongly reduced. Towards the channel center, similarly to the mean link-length $\overline{d^t}$, the transitivity differences for the three velocity components tend to reduce.

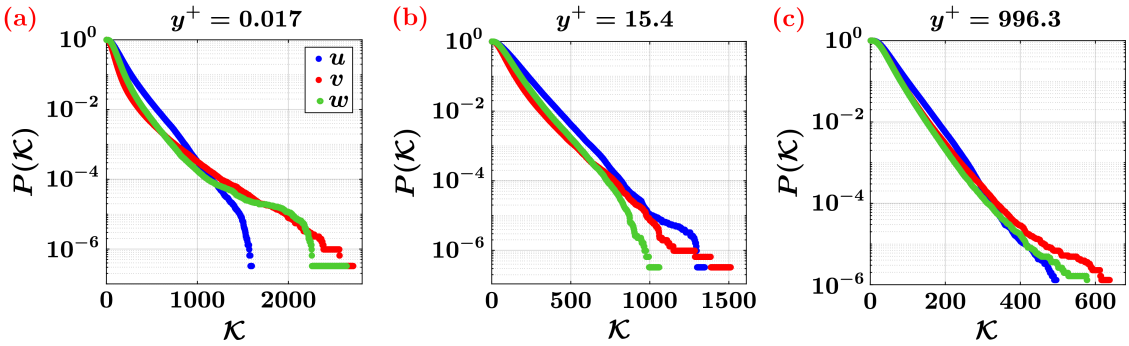


Figure 5.8: Cumulative degree distributions of the visibility networks built on the time-series of (u, v, w), and averaged over the homogeneous directions. (a): $y^+ = 0.0017$, (b): $y^+ = 15.4$, (c): $y^+ = 996.3$.

Finally, the cumulative degree distributions, $P(\mathcal{K})$, of the networks built on the three velocity components, (u, v, w), and averaged over the grid-points in the homogeneous directions are displayed in Figure 5.8. At fixed positions from the wall (i.e, $y^+ = \{0.017, 15.4, 996.3\}$), the slope of the three components is pretty similar, confirming that a steeper decay is present when moving far from the wall (from $y^+ = 0.017$ to $y^+ = 996.3$). However, it is worth to note that, very close to the wall (Figure 5.8(a)), the probability to find large values of degree, $\mathcal{K}$, is higher for the v and w components than for the u, despite the average value $\overline{\mathcal{K}}$ is the highest for u (see Figure 5.7(d)). Recalling that close to the wall the transversal components are more affected by the presence of extreme events (as highlighted by the mean link-length and assortativity in Figure 5.7(a),(c)), high degree values are more likely to appear for v and w. However, smaller values of $\mathcal{K}$ as less probable to appear (see Figure 5.8(a)), especially in the range $\mathcal{K} \in (100, 800)$, which substantially affects the average value, $\overline{\mathcal{K}}$.

5.3.3 Additional results

Here, further results on the investigation of velocity time-series in a turbulent channel flow via visibility networks are reported. First, a sensitivity analysis involving the time step of the network metrics is investigated in Section 5.3.1. Later,

a complementary study on the application of the horizontal visibility algorithm to the streamwise velocity time-series at two different Reynolds number is provided.

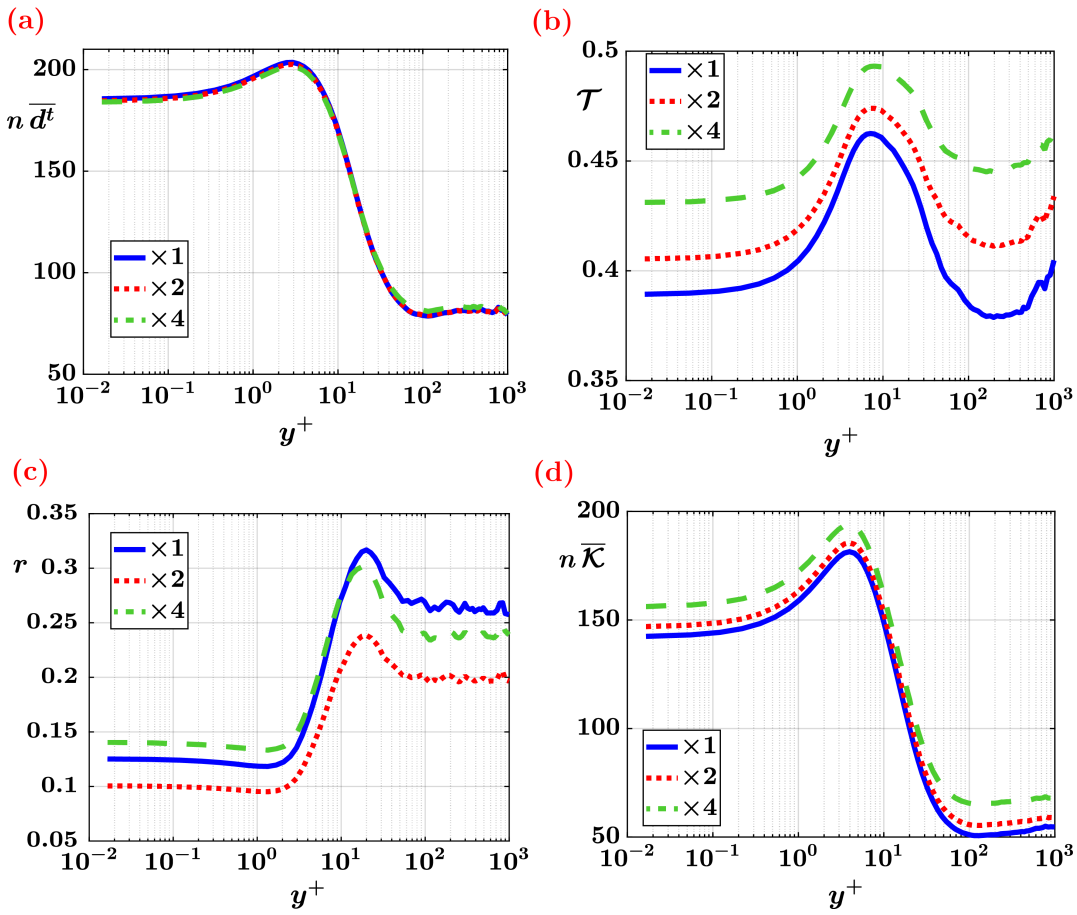


Figure 5.9: Averaged metric behaviours as function of y^+ for three different time sampling of the streamwise velocity time-series, u. The curves are obtained with $n\Delta t$, where the sampling is $n = 1$ (blue), $n = 2$ (red), $n = 4$ (green).

Parametric analysis on the time step

A parametric analysis on the temporal discretization is reported. A change in the number of time steps directly affects the number of nodes $\mathcal{N}_v$, and consequently the structure of the corresponding network. We recall that in the main analysis (Sections 5.3.1 and 5.3.2) the number of nodes was $\mathcal{N}_v = 4000$, with $\Delta t = 0.0065$ (i.e., $\Delta t^+ = 0.3246$). Two other time steps, namely $2\Delta t$, and $4\Delta t$, are here considered, resulting in networks with $\mathcal{N}_v = 2000$ and $\mathcal{N}_v = 1000$, respectively. In Figure 5.9, the metrics as function of y^+ are displayed for the three temporal samplings, $n\Delta t$, with $n = \{1, 2, 4\}$. Mean link-length and degree centrality are reported

as scaled with n (namely $n\overline{d^t}$ and $n\overline{\mathcal{K}}$), to facilitate the comparison between samplings. The transitivity, $\mathcal{T}$, and the assortativity coefficient, r, instead, are not scaled with n as by definition vary between 0 and 1. It can be observed that, apart from the specific values reached by the transitivity and the assortativity coefficient, the metrics behaviour along the wall-normal direction y^+ is not sensitive to the choice of the temporal discretization (i.e., the number of nodes). In other words, while the specific values of the metrics may change with Δt, the wall-normal variation of the network is preserved for the range of Δt investigated. This implies that, for example, the maximum values of the metrics occur at the same y^+ coordinates as Δt is changed. Accordingly, the behaviour of the metrics discussed in Sections 5.3.1 and 5.3.2 are insensitive to the specific time discretization adopted, considering that Δt is sufficiently small to preserve the temporal structure of the signal. Therefore, the visibility graph approach reveals to be a robust method to map time-series in networks, provided that the temporal recording is sufficiently long to ensure that the time-series are statistically stationary.

Horizontal visibility networks at different Re_τ

To enlarge and conclude the visibility graph analysis of turbulent channel flow, the *horizontal* visibility graph – which is the main variant of the natural visibility algorithm (see also Section 4.2.1) – is employed to map the streamwise velocity time-series. We recall that the horizontal visibility algorithm satisfies an ordering criterion (see Eq. (4.2)), so two data, $s(t_i)$ and $s(t_j)$, are linked if any other intermediate data, $s(t_k)$ with $i < k < j$, is smaller than both $s(t_i)$ and $s(t_j)$. For instance, Figure 5.10(a) shows an example of time-series (the same reported in Figure 5.2(a)), while the corresponding horizontal visibility graph is displayed in Figure 5.10(b). By comparing the two visibility variants, it can be observed that a lower number of links are activated by the horizontal visibility algorithm (i.e., links $4-6$, $4-9$ and $5-7$ are not active), as the ordering criterion is more restrictive than the convexity criterion. This implies that, in general, the overall connectivity in the networks (i.e., the node degree) is much lower for the horizontal visibility graphs. In particular, slow variations in the signal (e.g., as in the interval $4-7$ in Figure 5.10(a)) affect the overall connectivity of the horizontal visibility much more than the natural visibility, resulting in a much lower overall connectivity in the former case.

Here, three network metrics are focused: the mean link-length, the average clustering coefficient and the degree centrality. The average clustering coefficient, $\overline{\mathcal{C}}$ (see Eq. (3.7)), is selected instead of the transitivity because the latter is much more affected by the overall lower connectivity in the horizontal visibility graphs than the clustering coefficient. However, both the metrics are indices of the presence of triangles in the network, so $\overline{\mathcal{C}}$ is associated to the presence of irregularities in the signal. The behaviour of the selected metrics as a function of y^+ is investigated

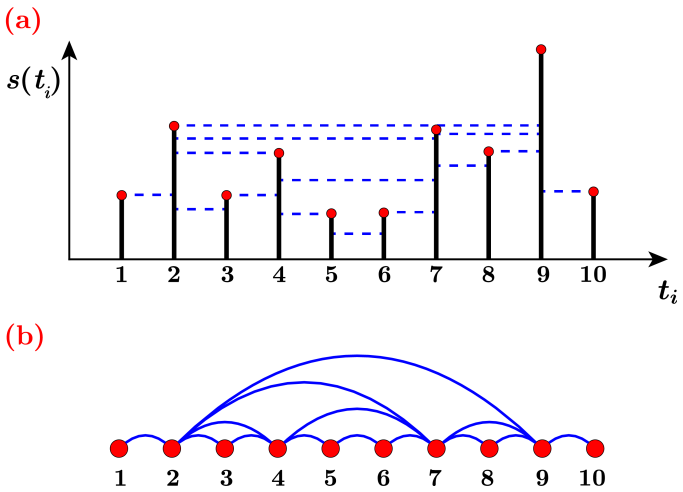


Figure 5.10: (a) Example of time-series and (b) network representation of the corresponding horizontal visibility graph. In (a) nodes are depicted in red, while visibility lines are illustrated as dashed blue lines.

for the time-series of u in two turbulent channel flows at $Re_\tau = 1000$ (i.e., the same dataset used in the main analysis of Sections 5.3.1 and 5.3.2) and $Re_\tau = 180$. Further details on the simulation at $Re_\tau = 180$ can be found in Appendix B.3.

Figure 5.11 shows the wall-normal behaviour of the network metrics. In general, all the metrics for the highest Reynolds number are larger than the metrics for $Re_\tau = 180$. Moreover, the minimum/maximum values of the metrics occur at close y^+ coordinates for the two Reynolds numbers. These results suggest that the visibility networks are able to capture the near wall dynamics at different Reynolds numbers, although the specific values are scaled for different Re_τ. More in detail on the metrics, the (average) mean link-length (multiplied by Δt^+ for comparison in Figure 5.11(a)) follows a trend as a function of y^+ that is similar to that for the natural visibility graph (see Figure 5.4(a)), for both the Reynolds numbers. Since $\overline{d^t}$ is a measure of the horizontal separation in the series, the results from the horizontal visibility are in accordance with those for the natural visibility. It should be noted that, while for the natural visibility approach only the temporal (horizontal) length of each link is considered in $\overline{d^t}$, for the horizontal visibility approach the mean link-length accounts for the total length of each links (which are intrinsically horizontal). This conceptual difference, therefore, can lead to variations in the $\overline{d^t}$ behaviour between the natural and horizontal visibility approaches.

The degree centrality, $\overline{\mathcal{K}}$, and the clustering coefficient, $\overline{\mathcal{C}}$, (Figure 5.11(b),(c)), instead, follows a wall-normal behaviour that is inverse with respect to the natural visibility (see Figure 5.4(b),(d)). The reason for the inverse behaviour of $\overline{\mathcal{K}}$ and $\overline{\mathcal{C}}$ with respect to the natural visibility is due to the fact that the link activation for the horizontal visibility is much more sensitive to slow variation in the signal.

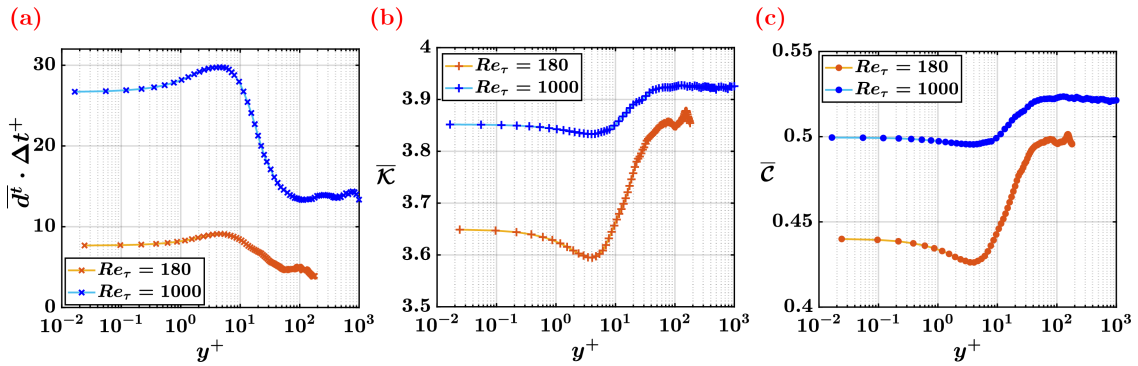


Figure 5.11: Vertical profiles of the average values of (a) mean link-length (in wall-units), (b) degree centrality and (c) clustering coefficient from horizontal visibility, for channel flows at $Re_\tau = 180$ and $Re_\tau = 1000$.

For example, in Figure 5.10(a) node $i = 6$ is not linked to node $j = 4$ because the value of the series in correspondence of node $k = 5$, i.e. $s(t_5)$, is very close to the value of $s(t_6)$; in the natural visibility, instead, nodes $i = 6$ and $j = 4$ are linked (see Figure 5.2(a)). Therefore, for the horizontal visibility, differently from the natural visibility, a rapidly fluctuating series (such as signals extracted close to the channel center) produces higher levels of average degree and clustering coefficient. Since peaks are more frequent far from the wall (as suggested by Figure 5.11(a)), it is more likely that links are activated. On the other hand, close to the wall, the slow temporal variation in the signals of u prevent the horizontal visibility, so that $\overline{\mathcal{K}}$ decreases as $y^+ \to 0$. For the same reason, the clustering coefficient decreases towards the wall, since a low number of triangles are formed due to the low connectivity in the proximity of the wall. It is worth noting that the average degree for high y^+ approaches the (analytical) value $\overline{\mathcal{K}} = 4$, which is found in a horizontal visibility graph built from a random/uncorrelated time-series [122].

In summary, as for the natural visibility, the horizontal visibility is able to capture information on the temporal structure of the signals. Although analytical results are available, the horizontal visibility algorithm (due to its simplicity) retains less information than the natural visibility, which is then preferred for network-based investigations. Moreover, even though results shown are not scaled with any parameter, the analogies between the metrics behaviour at different Reynolds numbers are promising, thus suggesting a more detailed analysis in future works.

5.4 Time-series analysis of a passive scalar plume in a turbulent boundary layer

Besides the application to velocity data extracted from numerical simulations, the natural visibility algorithm is also employed to experimentally investigate the dispersion of a passive scalar plume in a turbulent boundary layer. Specifically, we address the same issue studied by *Fackrell and Robins* [77] and *Nironi et al.* [71] to investigate the spatio-temporal development of a passive scalar plume, emitted from an elevated point-source of varying size.

Experimental measurements are performed in a recirculating wind tunnel with rough wall, in which develops a turbulent boundary layer with a free-stream velocity $U_\infty = 4.94$ m/s and a thickness $\delta = 314$ mm. Due to its density similar to air, ethane (C_2H_6) is used as a passive tracer. A mixture of air and ethane is continuously ejected from a L-shaped tube of diameter, D_s, located at a wall-normal height $h_s/\delta \approx 0.24$. Figure 5.12 shows a sketch of the setup, and an example of passive scalar concentration time-series with the corresponding visibility network representation. Two source configurations are selected for two different diameters, $D_s/\delta = 9.55 \times 10^{-3}$ ($D_s = 3$ mm) and $D_s/\delta = 1.91 \times 10^{-2}$ ($D_s = 6$ mm). Throughout this section, the two source configurations are referred to as D3 and D6, respectively. Simultaneous measurements of longitudinal and vertical velocity, u, v and concentration, c, are performed by means of a hot-wire anemometer (HWA) and a fast Flame Ionization Detector (FID), respectively. All the time-series are normalized by characteristic velocity and concentration scales so that, in the following, velocity and concentration signals are intended to be adimensional. Full details on the setup and the measurement procedures are reported in Appendix C.

The acquisition time is set equal to $T = 180$ s with a samplifing frequency of 1000 Hz, resulting into 1.8×10^5 data (which is equal to the number of nodes, $\mathcal{N}_v$ in the visibility networks). Measurements are performed at different locations along the three Cartesian directions, (x, y, z). Specifically, data are recorded at $x/\delta = \{0.325, 0.650, 1.30, 2.60, 3.90\}$ in the streamwise direction. For each x/δ location, one-point measurements are taken along the vertical (i.e., at fixed z/δ) and transversal (i.e., at fixed y/δ) directions. Transversal profiles of concentration and velocity are obtained at $y = h_s$, and at spanwise locations ranging in the interval $z/\delta = [-0.6,0.6]$. On the other hand, vertical profiles are obtained at $z/\delta = 0$ at various wall-normal locations ranging in the interval $y/\delta = [0.096,0.828]$ (the limits depend on the estimated size of the plume at a given x/δ). It should be noted that the minimum coordinate along the wall-normal direction in which measurements are performed is $y = 30$ mm, being the surface roughness height equal to 20 mm. Furthermore, we refer to the *near field* and the *far field* as the streamwise locations closest and farthest from the source, respectively. Therefore, the near and far fields correspond to locations $x/\delta = 0.325$ and $x/\delta = 3.90$, respectively, while $x/\delta = 1.30$

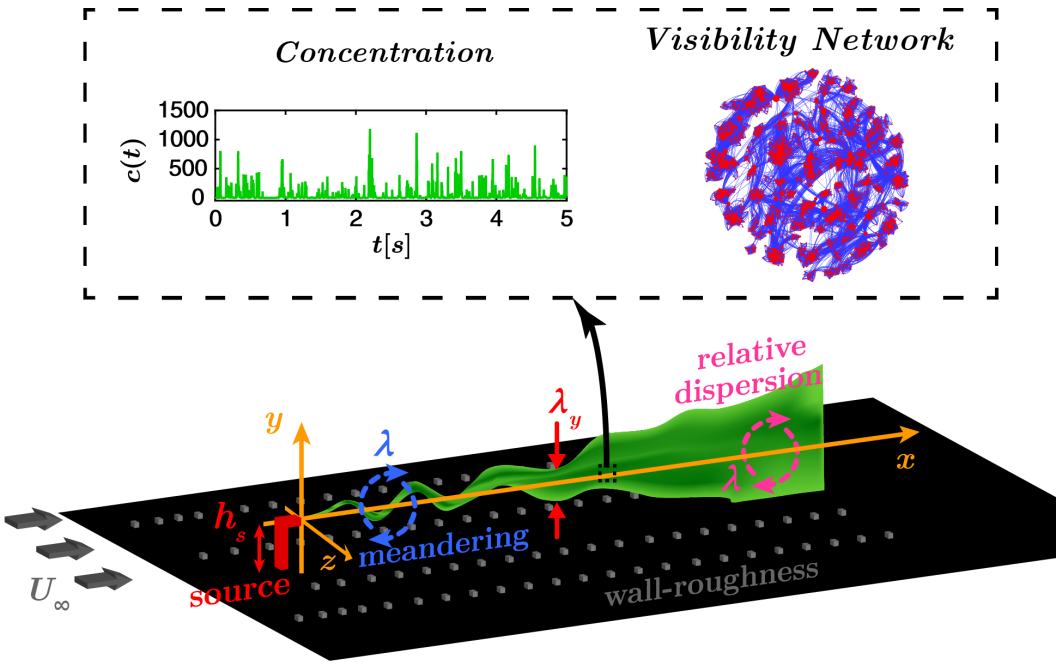


Figure 5.12: 3D sketch of the TBL setup in a (x-y) plane. The plume is illustrated in green, while the horizontal dash-dot line refers to the source axis. A zoomed view is shown within a dashed box containing the first 5 seconds of a concentration signal and the corresponding natural visibility network. The symbols are defined in the main text. Two large-scale eddies of (Eulerian) characteristic size λ are also depicted as rotating arrows.

is considered as an intermediate location.

In this work, we mainly focus on two quantities to characterize the plume dynamics: (i) the concentration field, c, and (ii) the vertical turbulent transport, $v'c'$. The principal (classical) statistics of concentration and vertical turbulent transport are discussed in Section 5.4.1. Vertical profiles of the network metrics extracted from the concentration and vertical turbulent transport time-series are then reported in Section 5.4.2 and Section 5.4.3, respectively. The vertical turbulent transport is here investigated as it plays a key role into the interplay between the turbulence velocity field and the concentration of passive scalar. Moreover, due to the crucial role played by the wall-normal direction that represents the direction of spatial inhomogeneity of the flow, the main analysis is focused on measurements taken at $z/\delta = 0$, namely in the (x, y)-plane normal to the wall and passing through the source axis.

Complementary results on the transversal profiles of concentration and vertical profiles of longitudinal turbulent transport are reported in Section 5.4.4. In order to obtain turbulent transport signals, the Reynolds decomposition is performed for

velocity and concentration time-series as $u' = u - U$, $v' = v - \overline{v}$ and $c' = c - \overline{c}$, where $\overline{v}$ and $\overline{c}$ are the (one point) time-averages of v and c, respectively. The effect of the source size, D_s, on the statistics and the network metrics is emphasized throughout the following sections of this chapter.

5.4.1 Wall-normal statistics of concentration and vertical turbulent transport

Previous works (e.g., Ref.s [71, 77]) focused on the influence of the source size on the one-point concentration statistics, at varying distance from the source. In a similar way, in this section, classical statistics of concentration, c, and vertical turbulent transport, $v'c'$, are shown and their behaviour along the wall-normal direction, y, is discussed. Specifically, lower order statistics of concentration (i.e., mean and standard deviation) are given as a comparison with the current literature; for the vertical turbulent transport, instead, higher-order statistics are also provided in order to enrich the benchmark for a dispersing plume in a rough-wall setup. Indeed, since turbulent transport requires simultaneous measurements of velocity and concentration, statistics on turbulent fluxes are not always available.

Mean and standard deviation profiles

The vertical profiles of mean and standard deviation of the (normalized) concentration are shown in Figure 5.13. The mean concentration profiles are rather similar for the two source diameters, D3 and D6, except in the proximity of the source where the plume statistics are affected by the presence of the source. In fact, the plume sizes – i.e. the transversal and wall-normal spread of the plume – become much more larger than D_s with increasing x/δ due to the relative dispersion (e.g., see Figure 5.12); as a result, the effect of D_s on $\overline{c}$ becomes negligible for $x/\delta \gg 0$. The difference in the standard deviation profiles between D3 and D6 is more pronounced than for the mean concentration, as an effect of the meandering of the plume in the proximity of the source. These results are in accordance with previous studies [71, 77], which show that, while the mean concentration profiles are almost unaffected by the source conditions, the higher-order statistics show a high sensitivity on the source size, even at large distances from the release point.
The vertical position, h_s^*, of the actual axis of the plume is also illustrated in Figure 5.13, for both D3 and D6; h_s^* is defined as the y coordinate of maximum $\overline{c}(y)$ value (see Appendix C, Section C.3 for more details on the evaluation of h_s^*). In fact, since the plume develops in a turbulent boundary layer, it is affected by the mean shear and by the source wake. While the latter is mainly present very close to the source, the mean shear acts at any streamwise location and tends to tilt the plume axis towards the wall. As a consequence, the wall-normal coordinate of the plume axis, h_s^*, is not exactly at the source axis, $y = h_s$, but h_s^* decreases

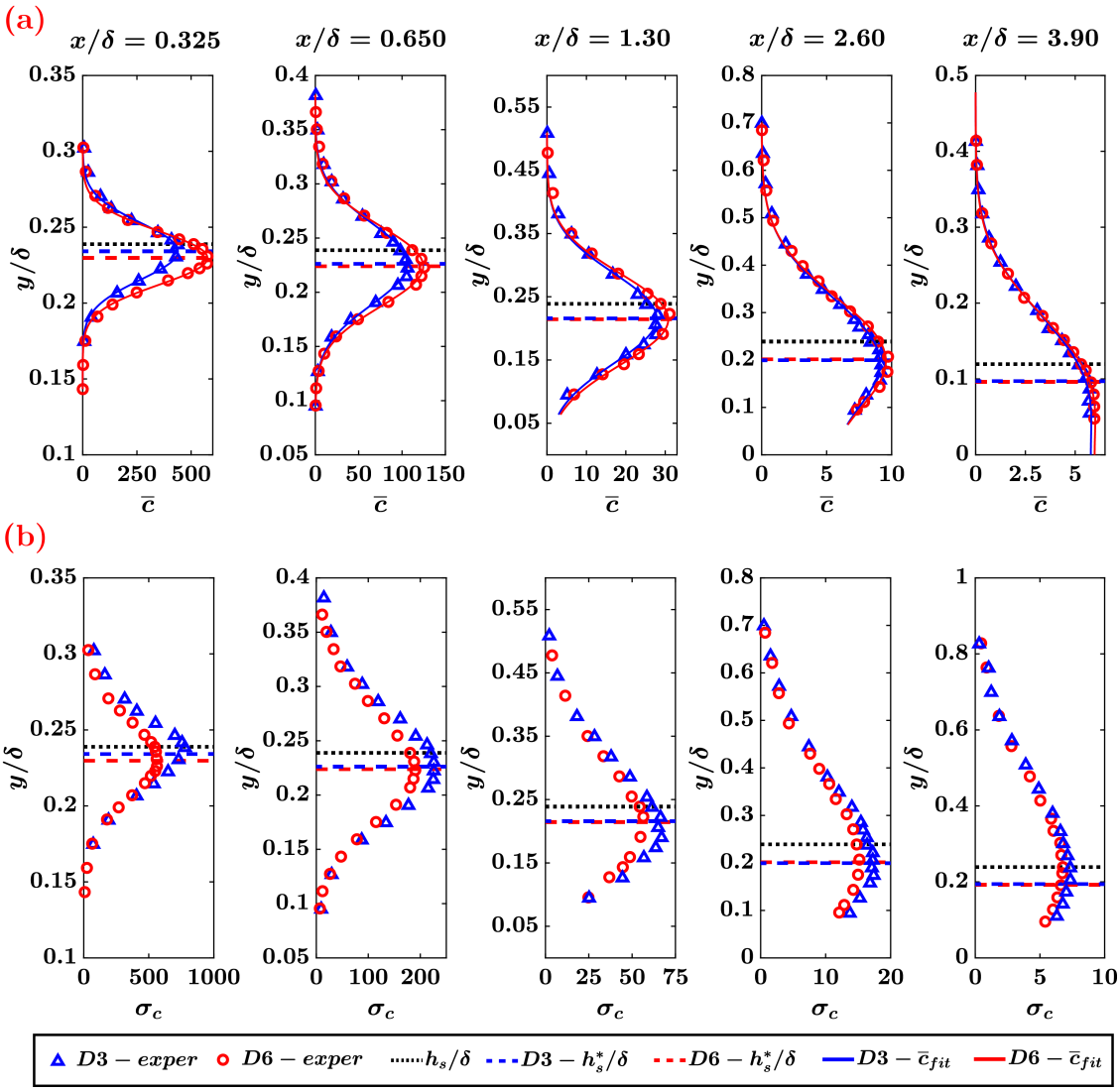


Figure 5.13: Vertical profiles of the mean and standard deviation values of the concentration, $\overline{c}$ and σ_c, at different streamwise locations, for the two source diameters, $D = 3$ mm and $D = 6$ mm. The profiles obtained from the reflected Gaussian distribution of Eq. (C.2) are also shown. The wall-normal coordinate of the source axis, h_s, is illustrated as a horizontal dotted line, while the plume axis height, h_s^*, is displayed as a blue (red) dashed line for the source D3 (D6).

downstream. Although the values of h_s^* for D3 and D6 should be different, this is true only in the near field, i.e. where the differences between the plumes emitted by D3 and D6 are the strongest.

Similarly to the concentration field, we examine here the statistics of vertical turbulent transport, $v'c'$. Figure 5.14 shows the vertical profiles of mean and standard

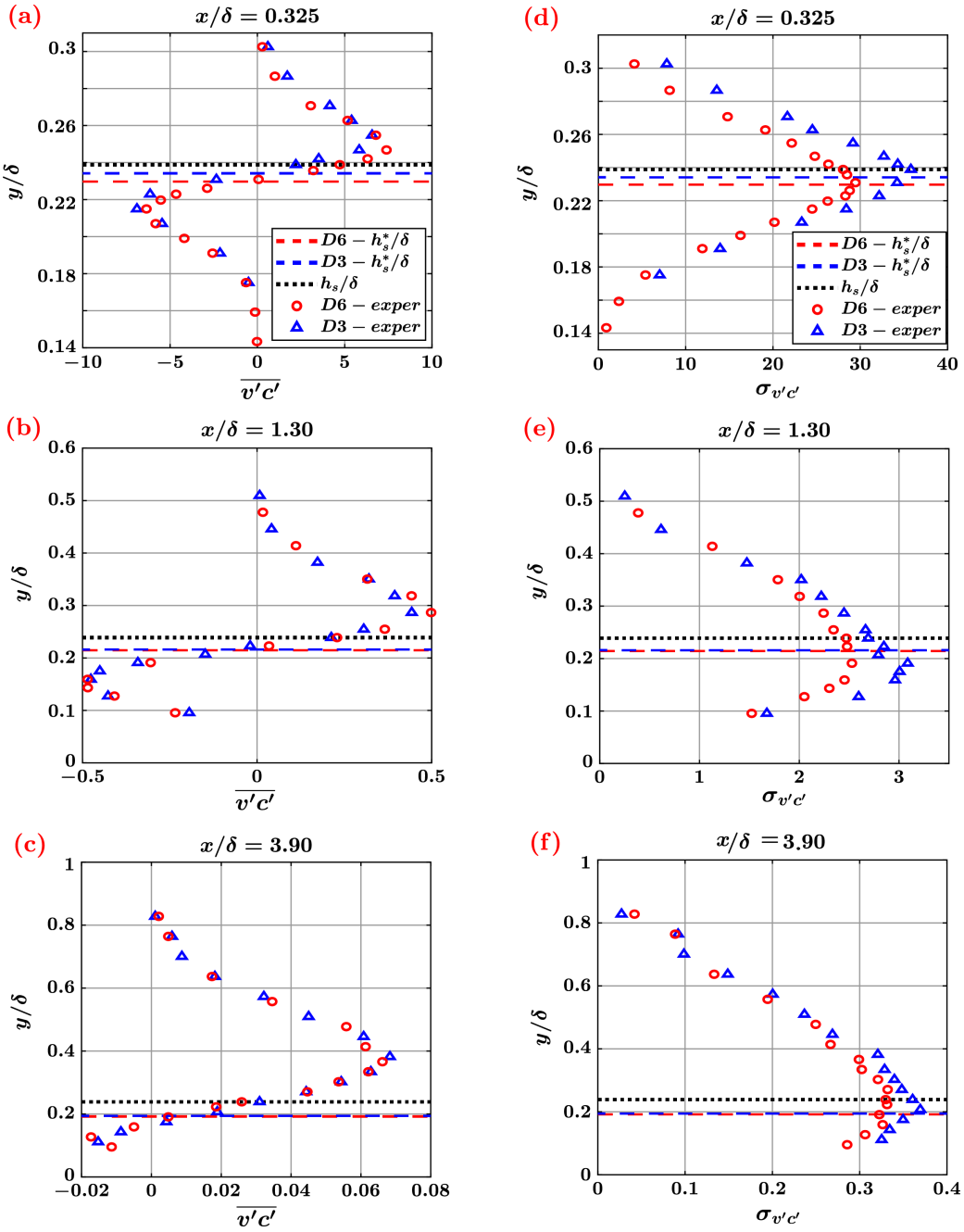


Figure 5.14: Vertical profiles of the mean value of wall-normal transport $\overline{v'c'}$ (a)-(c), and the standard deviation $\sigma_{v'c'}$ (d)-(f). The profiles are plotted at $x/\delta = \{0.325, 1.30, 3.90\}$, for the two source diameters, D3 and D6. The source axis height, h_s, is illustrated as a horizontal dotted line, while the plume axis height, h_s^*, is displayed as a blue (red) dashed line for the source D3 (D6).

deviation values of the (normalized) turbulent flux, $v'c'$. The profiles are reported at three representative streamwise locations, i.e. in the near field ($x/\delta = 0.325$), in the far field ($x/\delta = 3.90$), and at an intermediate location ($x/\delta = 1.30$). The vertical profiles of $\overline{v'c'}$ – namely the total mass transport – tend to collapse for the two configurations D3 and D6, as shown in Figure 5.14(a)-(c). This behaviour is more evident in the far field than in the proximity of the source, as the dependence of the mean concentration on the source size D_s rapidly vanishes downstream from the source. The vertical turbulent transport is zero at the plume axis, h_s^* (as an effect of the mean shear), namely where the vertical concentration gradient is minimum (see Figure 5.13(a)). Above and below the plume axis, instead, the mean value of turbulent flux is non-zero (see Figure 5.14(a)-(c)): for $y > h_s^*$, $\overline{v'c'}$ is positive while, for $y < h_s^*$, $\overline{v'c'}$ is negative. Above the source axis, the passive scalar is mainly carried out upwards by positive v' fluctuations; on the other hand, below the source axis the passive scalar is mainly transported downwards by negative v' fluctuations. In particular, the maximum/minimum value of $\overline{v'c'}$ corresponds to the maximum of the mean concentration gradient (this is also evident by using the Boussinesq approximation $\overline{v'c'} \sim -\partial\overline{c}/\partial z$ [181]).

As shown in Figure 5.14(d)-(f), the effect of the source size for an elevated source is instead much more evident for the standard deviation, $\sigma_{v'c'}$, rather than for the mean values – in analogy with what is observed in the concentration statistics – even at large distances from the source. This is a consequence of the stronger meandering motion of the plume emitted by the smallest source size, D3, which produces more variability in the series and the corresponding high intermittency in the dynamics of vertical turbulent transport (see also the discussion in Section 2.3.2). The maximum difference of standard deviation between D3 and D6 is present close to the plume axis in the near field, and such difference strongly decreases by moving downstream towards the far field due to the weakening of the meandering and the strengthening of the relative dispersion. By moving in the wall normal direction, $\sigma_{v'c'}$ decreases as the plume intensity vanishes away from the source axis.

Higher-order moments of vertical turbulent transport

The behaviour of the higher-order moments of turbulent transport, $v'c'$, is here discussed by focusing on the skewness, $Sk_{v'c'}$, and the kurtosis, $Ku_{v'c'}$. Formally, they are defined as the normalized third- and fourth-order central moments, namely $Sk_{v'c'} = \overline{(v'c' - \overline{v'c'})^3}/\sigma_{v'c'}^3$ and $Ku_{v'c'} = \overline{(v'c' - \overline{v'c'})^4}/\sigma_{v'c'}^4$, respectively. Figure 5.15 shows the skewness and the kurtosis as a function of y/δ for three streamwise locations (as for the mean and standard deviation shown in Figure 5.14). The behaviour of the skewness (Figure 5.15(a)-(c)) is similar for the two source configurations D3 and D6 at any x/δ. In particular, $Sk_{v'c'} \approx 0$ at the plume axis, while the skewness is negative/positive below/above the plume axis, because below and above h_s^* the vertical turbulent transport is mainly downwards (i.e., $v'c' < 0$) and

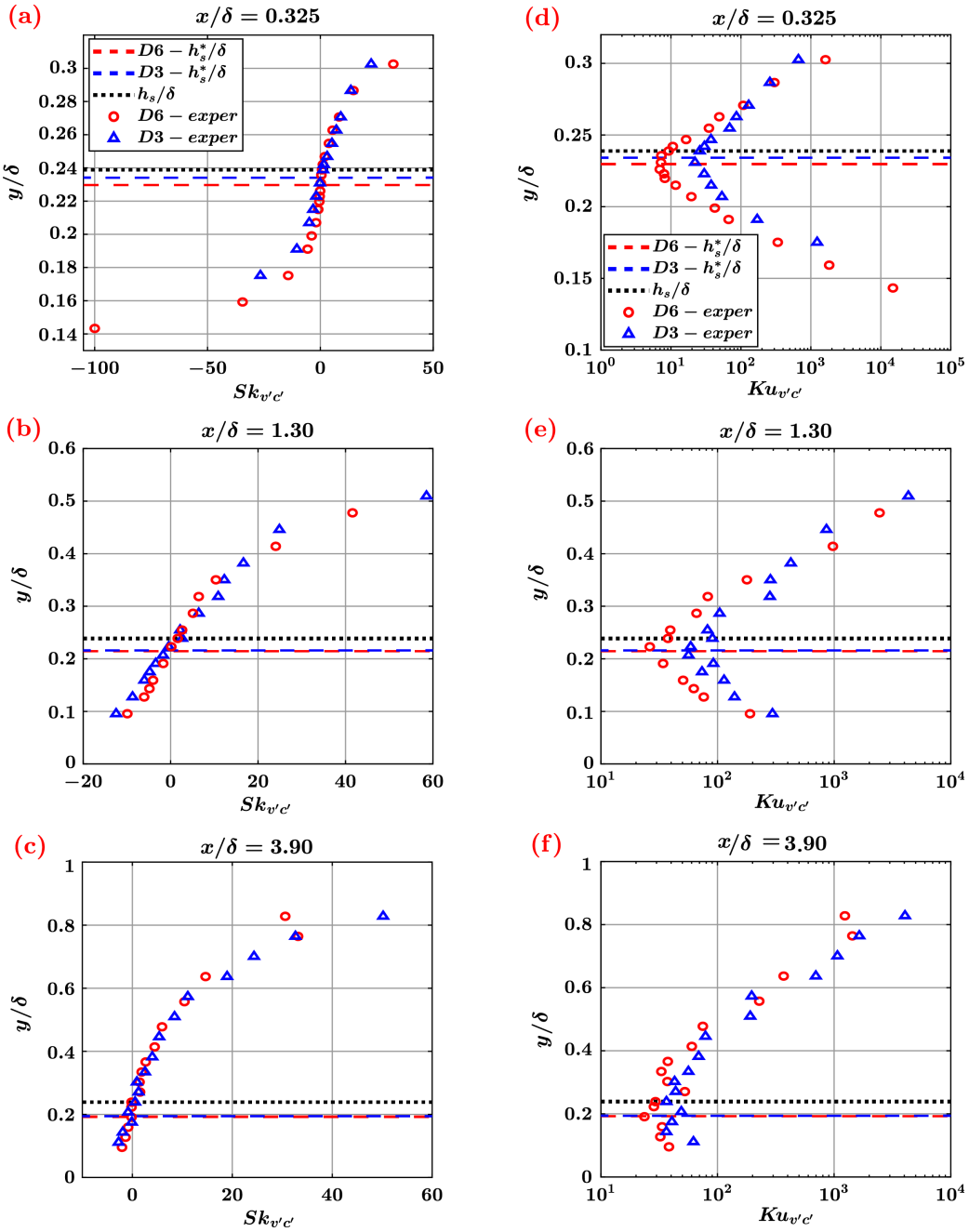


Figure 5.15: Vertical profiles of the skewness of wall-normal transport $Sk_{v'c'}$ (a)-(c), and the kurtosis $Ku_{v'c'}$ (d)-(f), in log-lin scale. The profiles are plotted at $x/\delta = \{0.325, 1.30, 3.90\}$, for the two source diameters, D3 and D6. The source axis height, h_s, is illustrated as a horizontal dotted line, while the plume axis height, h_s^*, is displayed as a blue (red) dashed line for the source D3 (D6).

upwards (i.e., $v'c' > 0$), respectively.
As shown in Figure 5.15(d)-(f), the kurtosis values are greater than three (which corresponds to normal distribution), thus implying that the PDFs of $v'c'$ are fat-tailed distributions (under some circumstances, the PDFs can be well fitted by a gamma distribution [71]). In particular, $Ku_{v'c'}$ is minimum at the plume axis at each streamwise location, as the plume develops around $y = h_s^*$ and extreme events (with respect to $\overline{v'c'}$) are less probable to appear; on the contrary, away from the plume axis extreme $v'c'$ values are more probable, as the signals are much more intermittent. Differently from the skewness, the kurtosis profiles for D3 and D6 are different in the near field (see Figure 5.15(d)), and the difference of $Ku_{v'c'}$ progressively reduces towards the far field (see Figure 5.15(f)). This implies that the meandering affects the behaviour of the standard deviation but also the behaviour of the kurtosis: the values of $Ku_{v'c'}$ for D3, in fact, are higher than the values of $Ku_{v'c'}$ for D6, namely extreme values are more probable for D3 than for D6.

Spectra and intermittency factor of vertical turbulent transport

Figure 5.16 shows the normalized power spectral density, $\widetilde{E} = E\delta/\sigma_{v'c'}$, of the signals $v'c'$, as a function of the normalized wavenumber $\widetilde{k} = 2\pi\delta/\lambda$, where $E(\widetilde{k})$ is the power spectral density, $\lambda = \overline{u}/f$ is a characteristic turbulent length scale (see the sketch in Figure 5.12), f is the frequency and $\overline{u}$ is the (local) mean streamwise velocity. Spectra are plotted along the source axis, namely for $y = h_s$. Since the instantaneous plume size, $\lambda_{y,z}$, depends on the source size and the spatial location, the relation between $\lambda_{y,z}$ and turbulent length scales, λ, affects the behaviour of the spectra. Here we focus on the wall-normal direction, i.e., on λ_y, but a similar reasoning holds for λ_z. In particular, the instantaneous plume size λ_y is generally smaller for D3 than for D6 in the near field, due to the spatial proximity of the source; by moving downstream, instead, λ_y increases and the difference of λ_y between D3 and D6 diminishes.

In the near field (see Figure 5.16(a)), the difference of spectral density between D3 and D6 is larger at small wavenumbers than at high wavenumbers. In fact, turbulent length scales, λ, larger than the (instantaneous) plume size, λ_y, contribute to the (instantaneous) plume meandering motion in the wall-normal direction (e.g., see the sketch in Figure 5.12). Therefore, since in the near field $\lambda_{z,\mathrm{D3}} < \lambda_{z,\mathrm{D6}}$, the differences of $\widetilde{E}$ between D3 and D6 at low $\widetilde{k}$ are more evident, because the plume for D3 is affected by a wider range ($\lambda_y < \lambda < \lambda_{max}$, or equivalently, $\widetilde{k}_{min} < \widetilde{k} < 2\pi\delta/\lambda_y$) of turbulent scales. On the other hand, at high wavenumbers (namely, small turbulent length scales), the spectral density for the two source sizes tends to coincide, as turbulent scales $\lambda < \lambda_y$ only promote the dispersion of the plume. The large scale fluctuations – induced by a wider range of turbulent scale in the near field – progressively weaken towards the far field (see Figure 5.16(b),(c)), so that the intensity of spectral density decreases with x/δ and approaches the same

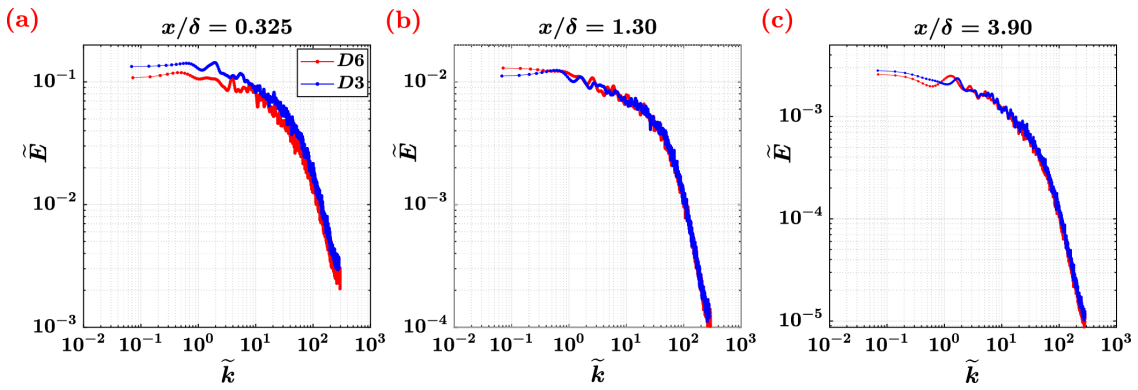


Figure 5.16: Normalized spectral density, $\widetilde{E}$, as a function of the normalized wavenumber, $\widetilde{k}$, of the wall-normal turbulent transport, $v'c'$. Spectra are evaluated at the source axis ($y = h_s$), for D3 and D6, at three streamwise locations: (a) in the near field ($x/\delta = 0.325$), (b) at an intermediate location ($x/\delta = 1.30$), and (c) in the far field ($x/\delta = 3.90$).

behaviour at all wavenumbers for D3 and D6. In fact, for increasing x/δ, the plume size increases (i.e., $\lambda_y \rightarrow \lambda_{max}$) and the range of scales for which $\lambda > \lambda_y$ decreases. In other terms, all turbulent scales tend to contribute to the relative dispersion of the plume in the far field. Finally, it should be noted that spectra of vertical turbulent transport normalized by its variance do not show a self-similar behaviour along the streamwise direction. This is in contrast to what has been recently reported for concentration series, which show a self-similar behaviour in the range $0.5 \leq x/\delta \leq 4$ [79].

The last parameter investigated is the intermittency factor. For the concentration series, an intermittency factor, γ_c, can be defined as the fraction of non-zero concentration values, where small γ_c values correspond to highly intermittent series [71]. In other words, γ_c is the fraction of time in which the passive scalar is measured. In a similar way, here we define the intermittency factor for the vertical turbulent transport as the fraction of time in which the passive scalar is transported upwards, $\gamma_+ = \text{prob}\,[v' > 0, c \neq 0]$, and downwards, $\gamma_- = \text{prob}\,[v' < 0, c \neq 0]$, with $(\gamma_+ + \gamma_-) = \gamma_c$ (by definition) and prob $[\bullet, \bullet]$ indicating the joint probability. In the definition of γ_+ and γ_-, the velocity fluctuations, v', impose the sign to the fluxes while the concentration discriminates between the presence ($c \neq 0$) or the absence ($c = 0$) of the plume. Therefore, although the overall intermittency is governed by the concentration field (i.e., $c \neq 0$ or $c = 0$), the velocity component introduces a directionality for the intermittency (namely $v' > 0$ or $v' < 0$). For example, the intermittency factor for the portion of the concentration signal shown in Figure 5.17(a) is $\gamma_c = 0.44$, because there are 22 non-zero values of $c(t_i)$ out of 50 values. Among the 22 non-zero values, 14 observations correspond to an upward motion (i.e., $\gamma_+ = 14/50 = 0.28$), while 8 observations correspond to a downward

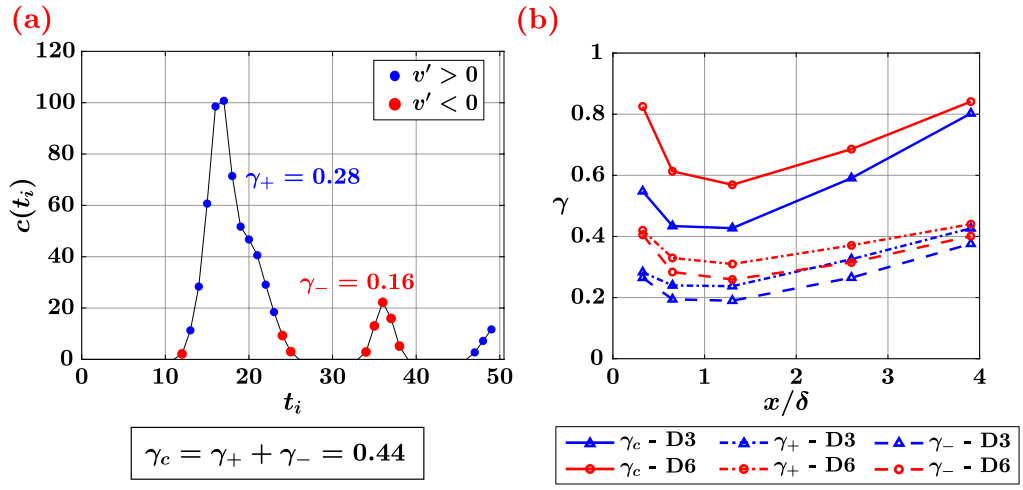


Figure 5.17: (a) Example of the intermittent behaviour of the concentration signal (first 50 values) measured at $x/\delta = 1.30$ and $y = h_s$. Blue and red data correspond to an upward and downward transport, respectively. The corresponding values of the intermittency factor for the shown time interval are also reported. (b) Intermittency factors of the passive scalar concentration, γ_c, and wall-normal turbulent flux, γ_+ and γ_-, as a function of x/δ along the source axis (i.e., $y/\delta = 0$ and $y = h_s$).

motion (i.e., $\gamma_+ = 8/50 = 0.16$).

Figure 5.17(b) shows that, at the source axis, the vertical transport is more intermittent downward ($v' < 0$) than upward ($v' > 0$), namely $\gamma_- < \gamma_+$ for both D3 and D6. Consistently with the results shown in Figure 5.14 and Figure 5.15, at the source axis the passive scalar is mainly transported upwards (as the plume axis lies below the source axis). Consequently, the fraction of time in which the passive scalar is transported upwards, γ_+, results to be greater than the fraction of time in which the passive scalar is transported downwards, γ_-. Furthermore, the intermittency factor is always smaller for the source diameter D3 than for D6, whether γ_c, γ_+ or γ_- is considered. This validates the fact that meandering motion is stronger for the plume emitted by a smaller source, inducing higher intermittency in the signals. More in detail, in the near field, the values of intermittency factor for D3 and D6 are different while they approach the same value in the far field, as the effect of the meandering is replaced by the relative dispersion of the plume. Although a strong meandering motion is present in the near field, the intermittency factors do not monotonically increase with x/δ because of the effect of the source proximity in the near field (as mentioned previously for the mean turbulent transport). A minimum value of the intermittency is found at $x/\delta \approx 1.3$, which is also found by *Nironi et al.* [71] in the case of γ_c.

Summary of the statistical analysis

The study of statistics reveals that the mean value and the skewness are not substantially affected by the source size, while the standard deviation and the kurtosis are sensitive to the emission conditions. The plume meandering – that is mainly active in the near field of the source – is the main responsible for the differences in the statistics: a stronger meandering motion produces higher variability (i.e., standard deviation) and more extreme events (i.e., kurtosis) for the smallest source size than for the largest source size. Far from the source, the meandering motion strongly reduces its intensity and the relative dispersion turns out to be the principal mechanism affecting the plume dynamics. In the far field, therefore, the statistics approach the same values for the two source sizes. The effect of the meandering on turbulent transport is also highlighted by the power spectral density and the intermittency factor of $v'c'$. Larger values of power spectral densities are obtained at small wavenumbers for the smallest source size, as a consequence of a wider range of turbulent scales affecting the plume dynamics. Additionally, the stronger meandering motion associated with the plume emitted by a smaller source size induces a more intense intermittency in the signals, namely a lower fraction of time for which the passive scalar is measured and transported. As for the statistical moments, the main differences of power spectral density and intermittency between the two source sizes are large in the source proximity and vanish in the far field.

5.4.2 Visibility-based investigation of concentration time-series

The results of the visibility network-based analysis applied to the passive scalar concentration time-series, c, are here reported and discussed. Results of both diameter configurations are shown in Figure 5.18. In particular, the mean link-length, $\overline{d^t}$ (Figure 5.18(a)-(b)), the transitivity, $\mathcal{T}$ (Figure 5.18(c)-(d)), and the assortativity coefficient, r (Figure 5.18(e)-(f)), are plotted as a function of the normalized wall normal coordinate $(y - h_s)/\delta$, for increasing streamwise locations. Moreover, in order to highlight the significance of the metrics with respect to the dynamics of the passive scalar plume, in Figure 5.19 we display two time-series of concentration (normalized with zero mean and unitary standard deviation) at two wall-normal coordinates, $(y - h_s)/\delta = 0$ (i.e., the source-axis) and $(y - h_s)/\delta = 0.27$.

First, the configuration with the smallest source size, $D_s = 3$ mm (i.e., D3) is focused. As illustrated in Figure 5.18(a), $\overline{d^t}$ increases by moving away from the source-axis in the wall-normal direction, and along the streamwise direction at fixed wall-normal coordinate. The low $\overline{d^t}$ values obtained along the source-axis suggest that peaks of the passive scalar concentration are frequently detected (e.g., see the series in Figure 5.19(a)), because along the source axis the effects of the relative dispersion of the plume are weak. On the other hand, the mean link-length

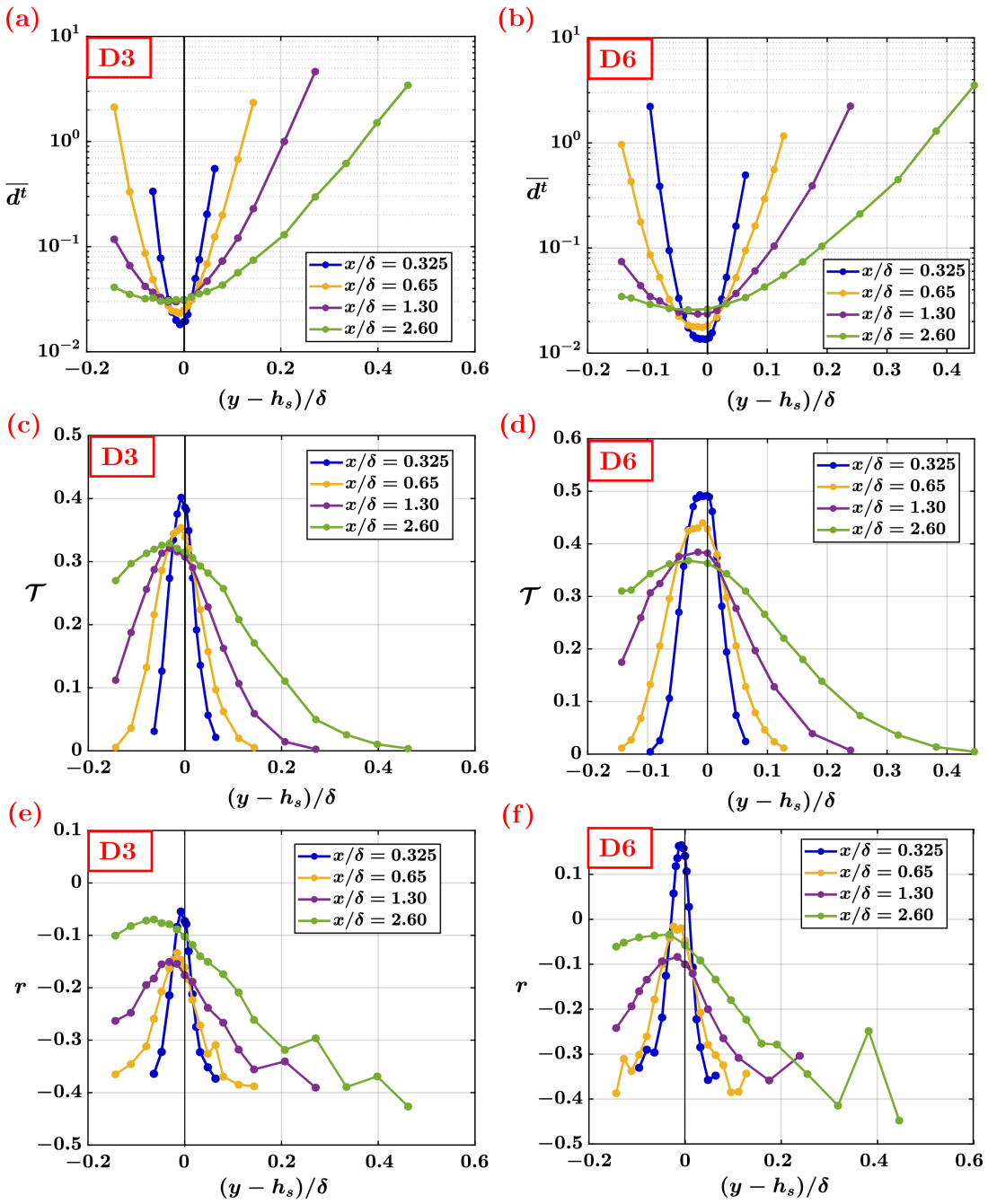


Figure 5.18: Network metrics as a function of the normalized wall-normal coordinate $(y - h_s)/\delta$, for different streamwise locations, x/δ, and for D3 (left panels) and D6 (right panels). (a)-(b) Mean link-length, (c)-(d) transitivity and (e)-(f) assortativity coefficient.

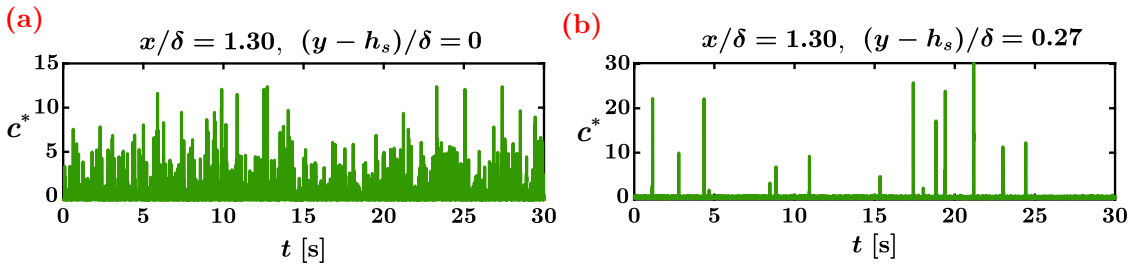


Figure 5.19: First 30 seconds of the time-series of ethane concentration at two different wall-normal coordinates for D3: (a), at the centreline $(y - h_s)/\delta = 0$; (b) $(y - h_s)/\delta = 0.271$. For comparison purposes, the time-series are normalized as $c^*(t) = (c(t) - \bar{c})/\sigma_c$.

increases in the wall-normal direction, because the occurrence of concentration measurements decreases along y as a consequence of the relative dispersion of the plume (e.g., see Figure 5.19(b)). Moreover, we found higher $\overline{d^t}$ values by moving downstream along x for wall-normal coordinates close to the source-axis, because the plume weakens along the streamwise direction around $y = h_s$. Conversely, we found lower $\overline{d^t}$ values by moving downstream far from the source axis, due to the increasing spatial extension of the plume caused by the relative dispersion. In fact, the average plume size, $\Lambda_{y,z}$, is larger in the far field than in the near field. Therefore, it is less likely to find outliers in the signal of c at a given wall-normal coordinate in the far field than in the near field (e.g., see in Figure 5.18(a) the difference of $\overline{d^t}$ between $x/\delta = 0.65$ and $x/\delta = 2.60$ for $(y - h_s)/\delta = 0.1$).
These outcomes are supported by the investigation of the transitivity, $\mathcal{T}$, and the assortativity coefficient, r. As shown in Figure 5.18(c) and Figure 5.18(e), the maximum values of $\mathcal{T}$ and r are found along the centreline of the plume. As discussed for the mean link-length, at the source axis it is more likely to find a high recurrence of non-zero values of ethane concentration (see Figure 5.19(a)), so that extreme events appear more as peaks than outliers. On the contrary, far apart from the source-axis the temporal structure of the series of concentration is dominated by the intermittency (i.e., long intervals of very low intensity fluctuations), thus producing very low values of transitivity and assortativity (i.e., the series are dominated by small intensity fluctuations with sporadic outliers). This behaviour can be observed, for example, in the time interval $t \in [11,15]$ and $t \in [24,30]$ in the series of Figure 5.19(b). Following specular trends with respect to $\overline{d^t}$, the transitivity and the assortativity coefficient decrease by moving away from the source-axis.

The metrics extracted from the network built on the concentration time-series in the configuration D6 are shown in Figure 5.18, panels (b),(d) and (f). All the three metrics highlight the same overall dynamics of the passive scalar plume discussed for D3. In fact, the minimum value of $\overline{d^t}$ and the maximum value of $\mathcal{T}$ and r are found along the source centreline, while the three metrics increase/decrease moving

away from the source-axis and along the streamwise direction, as for D3.

By comparing the magnitude of the metrics in the same spatial locations for D3 and D6, we found lower values of $\overline{d^t}$ and higher values of $\mathcal{T}$ and r in the configuration D6 with respect to D3. This means that, at a fixed spatial location, the concentration for D6 is characterized by more frequent non-zero measurements, namely a higher turbulent fluctuating activity is detected. This is due to the fact that the release emitted by the larger source is submitted to a reduced spatial range of velocity fluctuations that are able to displace the scalar plume as a whole (namely, the plume meandering). Therefore, a weaker intermittency is induced in the concentration signals (measured at fixed locations) for D6 than for D3. It is worth noting that – for both D_s configurations – the results are asymmetrical with respect to the source-axis due to the presence of the wall.

Finally, it is worth highlighting that the maximum/minimum values of the metrics are not found exactly at the source axis, $y = h_s$, but at a slightly lower wall-normal coordinate that is the plume axis, $y = h_s^*$. Furthermore, the values of assortativity coefficient, r, for the signals of concentration are almost always negative (see Figure 5.18(e),(f)), while for the velocity signals in the channel flow r is always positive (see Figure 5.7(c)). This reflects the substantial difference in the time-series structure between the concentration (which is more intermittent, see Figure 5.19) and velocity (which varies more regularly, e.g., see Figure 5.5).

5.4.3 Extreme turbulent transport events in the view of visibility networks

The results from the application of the natural visibility algorithm to the time-series of turbulent transport, $v'c'$, are reported in this section. The two configurations of source diameter, D3 and D6, are displayed for different downstream locations, x/δ, and at various wall-normal coordinates, y/δ. Differently from the analysis of concentration time-series, here the average peak occurrence, $\phi = 1/\overline{d^t}$, is employed in place of the mean link-length, $\overline{d^t}$, in order to highlight the spatial regions in which peaks frequently occur. Additionally, the assortativity coefficient, r, is also focused so that, by combining ϕ and r, a characterization of the temporal structure of $v'c'$ in terms of extreme events is provided.

The behaviours of the average peak occurrence and the assortativity coefficient are shown in Figure 5.20 and Figure 5.21, respectively, as a function of the vertical coordinate, y/δ. In general, ϕ and r have their maximum values close to the plume axis, h_s^* (see horizontal dashed lines in Figure 5.20 and Figure 5.21). This means that, similarly to the concentration field, peaks frequently occur around the plume axis; however, differently to the concentration field in which r is generally negative, the vertical separation between peaks and the other data in the series is weak (since $r > 0$). This behaviour of the metrics for $v'c'$ is due to the fact that the plume is mainly located around the source axis while it meanders and develops

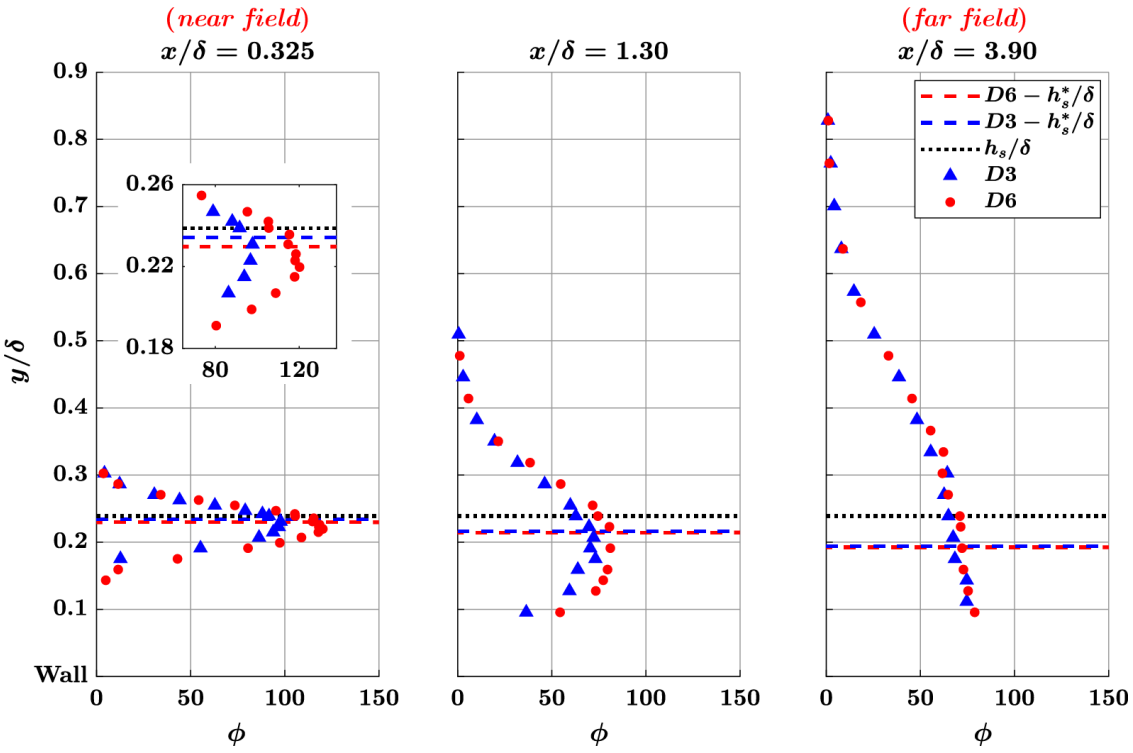


Figure 5.20: Vertical profiles of the average peak occurrence ϕ [Hz] based on signals of $v'c'$. The profiles are plotted in the near field and in the far field, for D3 (blue marker) and D6 (red marker). The source axis height, h_s, is illustrated as a horizontal dotted line, while the plume axis height, h_s^*, is displayed as a blue (red) dashed line for the source D3 (D6). The inset at $x/\delta = 0.325$ shows a zoom around the plume axis.

downstream. The difference in the features of temporal structure of the signals of $v'c'$ can be observed in Figure 5.22, which shows segments of time-series of $v'c'$ measured at $x/\delta = 0.325$ for D3 (Figure 5.22(a)-(c)) and for D6 (Figure 5.22(d)-(f)). Specifically, high values in the signals are much more frequent (i.e., high ϕ) and much less prominent (i.e., high r) around the plume axis (see Figure 5.22(b),(e)) than away from it.

Near field analysis. The effect of the source size, D_s, on the metrics is clearly visible in the near field (i.e., $x/\delta = 0.325$) and also at intermediate streamwise distances (i.e., $x/\delta = 1.30$) around the plume axis, where the metrics for D3 have smaller values than the metrics for D6. The peaks tend to appear more as outliers (i.e., lower r) and they occur less frequently (i.e., lower ϕ) for the smallest source size D3 (see Figure 5.22(b)) than for D6 (see Figure 5.22(e)). Since a plume emitted from a smaller source diameter is affected by a wider range of turbulent scales,

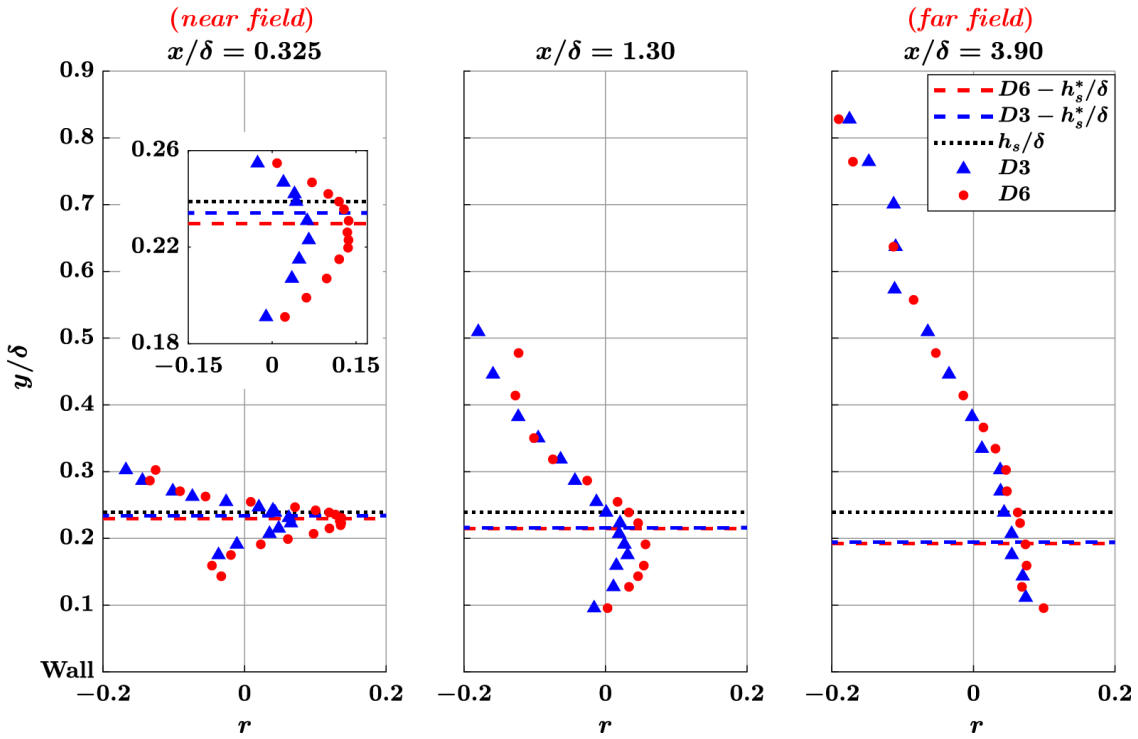


Figure 5.21: Vertical profiles of the assortativity coefficient r based on signals of $v'c'$. The profiles are plotted in the near field and in the far field, for D3 (blue marker) and D6 (red marker). The source axis height, h_s, is illustrated as a horizontal dotted line, while the plume axis height, h_s^*, is displayed as a blue (red) dashed line for the source D3 (D6). The inset at $x/\delta = 0.325$ shows a zoom around the plume axis.

its meandering motion is more intense. The unequal intensity of the meandering motion between D3 and D6 differently affects the temporal structure of turbulent transport series in the near field. Since a strong meandering motion induces a high intermittency in the series, it is expected that two peaks of turbulent transport are more unlikely to appear close in time (as there must be large fraction of time in which the passive scalar is not measured). Furthermore, the high variability (i.e., standard deviation) associated with the meandering motion implies that the relative intensity of peaks – with respect to the other values in the signal– increases, namely extreme events appear more as outliers because the time-series between two peaks is made up of values that are different from the peaks (otherwise the standard deviation values would be small). It should be noted that the networks corresponding to locations around the plume axis do not show large negative r values, as the higher variability in the signals prevent a strong separation between small values and peaks (which would give $r < 0$).

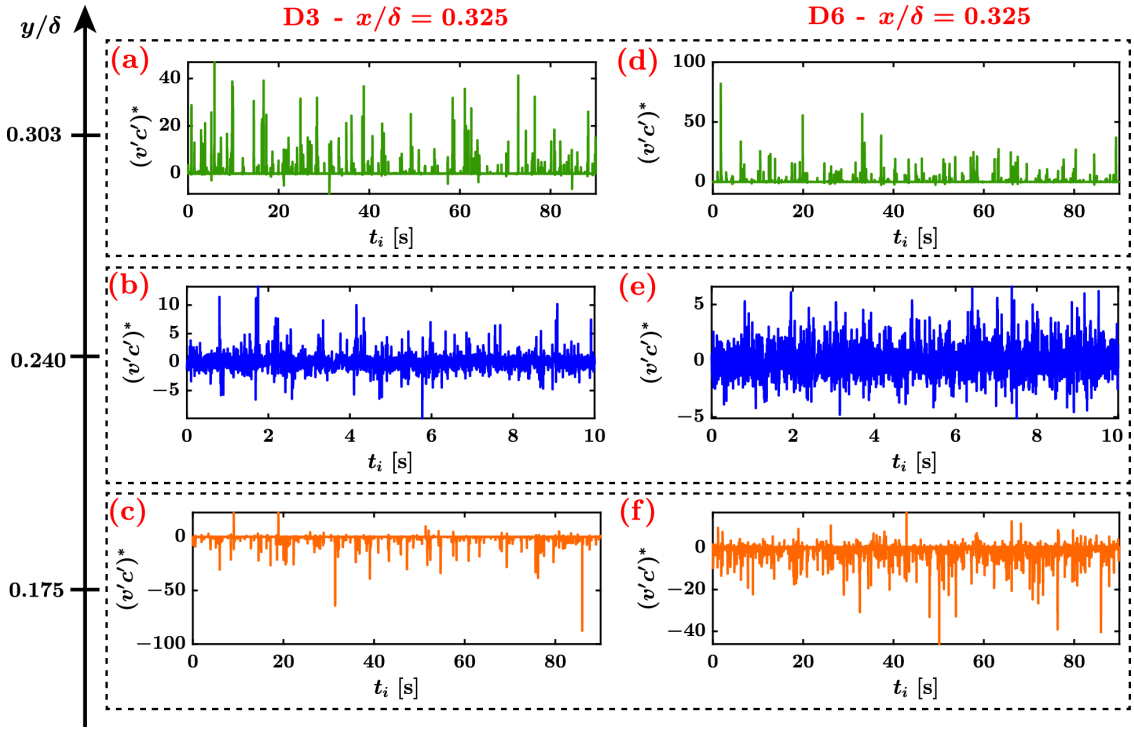


Figure 5.22: Time-series of vertical turbulent transport in the near field, $x/\delta = 0.325$, for the source diameter D3 (a)-(c) and D6 (d)-(f). Signals are plotted at three wall-normal coordinates, i.e., $y/\delta = 0.303$, $y/\delta = 0.240$ (the source axis), and $y/\delta = 0.175$. For comparison purposes, the time-series are normalized as $(v'c')^* = \left(v'c' - \overline{v'c'}\right)/\sigma_{v'c'}$.

Far field analysis. In the far field ($x/\delta = 3.90$), instead, the vertical profiles of the two metrics tend to collapse for both source diameter configurations, D3 and D6. This behaviour is a consequence of the increase of the plume size by moving downstream, due to the relative dispersion. In the far field the plume size exceeds the largest turbulent scales and the mixing of the passive scalar is then fully regulated by the relative dispersion rather than by the meandering. Therefore, by moving downstream around the plume axis, the average peak occurrence always decreases with x/δ, since large concentration values (and in turn high turbulent transport) are less probable to appear in the far field because of the plume weakening. On the other hand, for $y \approx h_s^*$, the assortativity coefficient first decreases (reaching a minimum at about $x/\delta = 1.30$) and then increases with x/δ. The behaviour of the assortativity coefficient is still a consequence of the interplay between the reducing meandering motion and increasing dispersion of the plume as it evolves downstream. It is worth noting that the minimum value of r along the source axis is found at $x/\delta \approx 1.3$, that is the streamwise location where the plume reaches the ground and all the intermittency factors are minimum (see also Figure 5.17(b)).

Wall-normal analysis. By focusing on the effect of the wall normal coordinate, y/δ, at a given streamwise location, the two metrics tend to decrease by moving away from the plume axis in the wall normal direction. Low values of ϕ and r indicate that, from the point of view of the temporal structures of the series, peaks occur less frequently and such peaks appear as outliers (e.g., see Figure 5.22(a),(d) and Figure 5.22(c),(f) for series above and below the plume axis, respectively). Therefore, the temporal structure of the signals becomes spike-like moving away from the plume axis along the y direction (see Figure 5.22(a),(c) for D3 and Figure 5.22(d),(f) for D6). It is worth noting that, similarly to the metrics for the concentration, in the near field the difference of average peak occurrence, ϕ, between D3 and D6 is larger below than above the plume axis (e.g., see the first-left panel of Figure 5.20). Therefore, the behaviour of ϕ in the near field indicates that peaks in the series of D6 are more frequent than for D3 below the plume axis. This behaviour can be seen, for instance, by looking at the signals shown in Figure 5.22(c),(f). Since in the near field the effect of the source size is still notable, the plume for D6 is larger and more tilted towards the wall than for D3 (i.e., h_s^* is smaller for D6). As a consequence, at a given vertical coordinate below the plume axis (e.g., at $y/\delta = 0.175$ in Figure 5.22) the temporal structure of the series for D6 (e.g., Figure 5.22(f)) is less spike-like than for D3 (e.g., Figure 5.22(c)).

Peak-pit asymmetry analysis. Finally, we remark that the visibility algorithm emphasizes the presence of positive peaks in the signals. However, as shown in Figure 5.22, there is an asymmetry between positive and negative extreme values in the series (especially away from the plume axis), as the passive scalar can be transported by turbulence upwards ($v' > 0$) and downwards ($v' < 0$). Therefore, the analysis of ϕ and r can also be carried out by applying the visibility algorithm to the signals $-v'c'$, thus highlighting the temporal occurrence and relative intensity of the negative extreme values (i.e., pits) of vertical turbulent transport.

The behaviour of ϕ and r obtained from the series $v'c'$ (top-visibility) and from the series $-v'c'$ (bottom-visibility) are shown in Figure 5.23 and Figure 5.24 for D3 and D6, respectively. The intersection between the metrics from the top- and bottom-visibility accurately corresponds to the vertical location of the actual plume axis, h_s^*, for both the source configurations D3 and D6. In fact, in the plume axis the vertical turbulent transport is expected to be symmetrical upwards and downwards (e.g., see Figure 5.22(b),(e)). On the other hand, the peak-pit asymmetry is significantly observed in the metric values above and below the plume axis, h_s^*. In particular, the metrics from the bottom-visibility are always smaller than the metrics from the top-visibility when the region below the plume axis is focused, while the opposite behaviour is found above the plume axis. The assortativity coefficient, r, for the bottom-visibility does not reach strong negative values for $y > h_s^*$, in both cases D3 and D6. This implies that there is not a strong vertical separation between pits and the other data values in the series of $-v'c'$, i.e. extreme negative values are

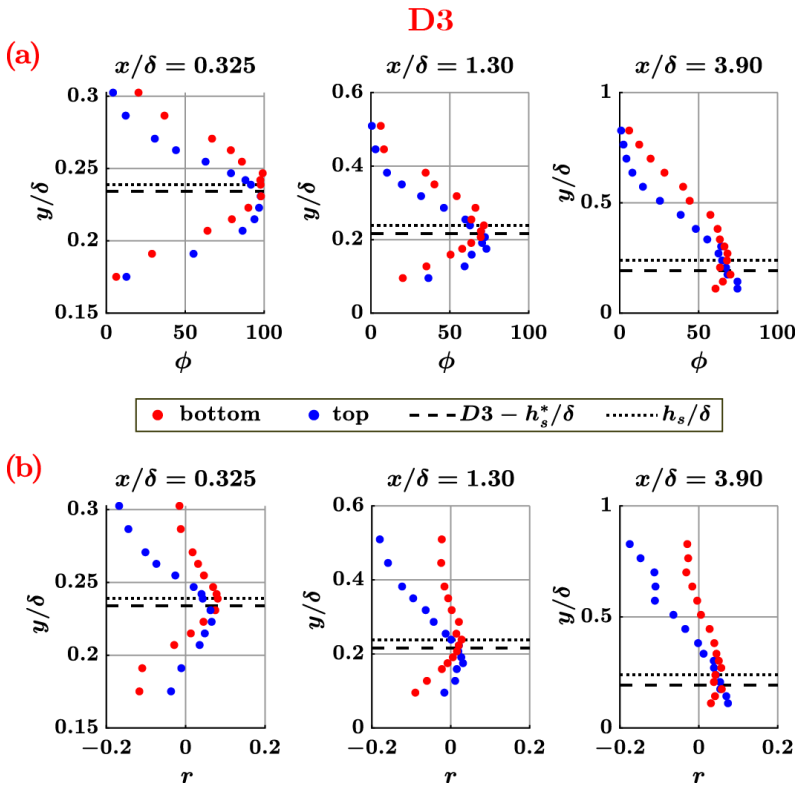


Figure 5.23: Vertical profiles of (a) the average peak occurrence, ϕ, and (b) the assortativity coefficient, r, for the top- and bottom-visibility in the configuration D3. The source axis height, h_s, is illustrated as a horizontal dotted line, while the plume axis height, h_s^*, is displayed as a dashed horizontal line for both D3 and D6.

unlikely to appear for $y > h_s^*$ (as shown in Figure 5.22(a),(d)). In fact, since most of the passive scalar is present around the plume axis, for $y > h_s^*$ vertical turbulent transport is mainly upwards while for $y < h_s^*$ it is mainly downwards.
In order to understand why the average peak occurrence, ϕ, for the bottom-visibility is larger than for the top-visibility for $y > h_s^*$, we recall that the visibility algorithm is insensitive to the absolute intensity of the data in the series. Therefore, above the plume axis, the visibility algorithm mostly considers pits in the series as negative values that are close to zero, because large negative values are very unlikely to appear. As a result, pits are found to frequently appear in time (i.e., high ϕ) and without a strong vertical separation (i.e., high r). On the contrary, the behaviour of the network metrics below the plume axis is the opposite with respect to the metrics above the plume axis, as the series above the plume axis (Figure 5.22(a),(d)) show an opposite temporal structure than the series below the plume axis (Figure 5.22(c),(f)).

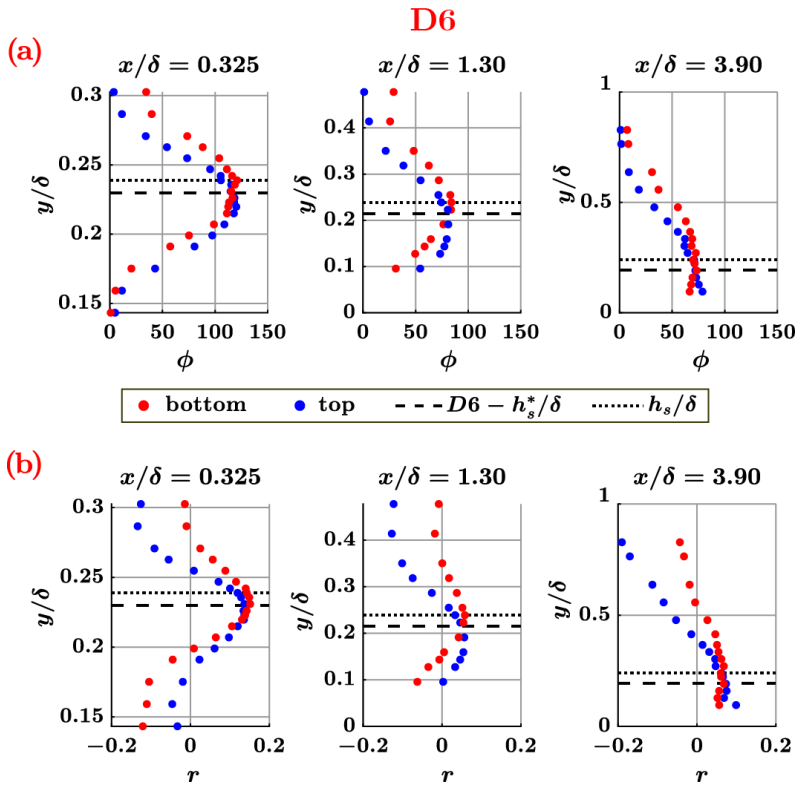


Figure 5.24: Vertical profiles of (a) the average peak occurrence, ϕ, and (b) the assortativity coefficient, r, for the top- and bottom-visibility in the configuration D6. The source axis height, h_s, is illustrated as a horizontal dotted line, while the plume axis height, h_s^*, is displayed as a dashed horizontal line for both D3 and D6.

To summarize, the respective influence of the meandering motion and the relative dispersion for different source sizes, as well as the plume weakening with the streamwise distance from the source, is fully captured by the network metrics. Specifically, the average peak occurrence and the assortativity coefficient are able to highlight the different temporal structures of the series – in terms of peaks occurrence and their relative intensity – at different spatial locations.

5.4.4 Complementary results

In this Section we report results that are complementary to the main analyses carried out in the previous Sections 5.4.2 and 5.4.3, for the concentration and vertical turbulent transport signals, respectively. In particular, the transversal profiles of the metrics for the concentration time-series, as well as the vertical profiles of the metrics for the longitudinal turbulent transport are reported and briefly discussed.

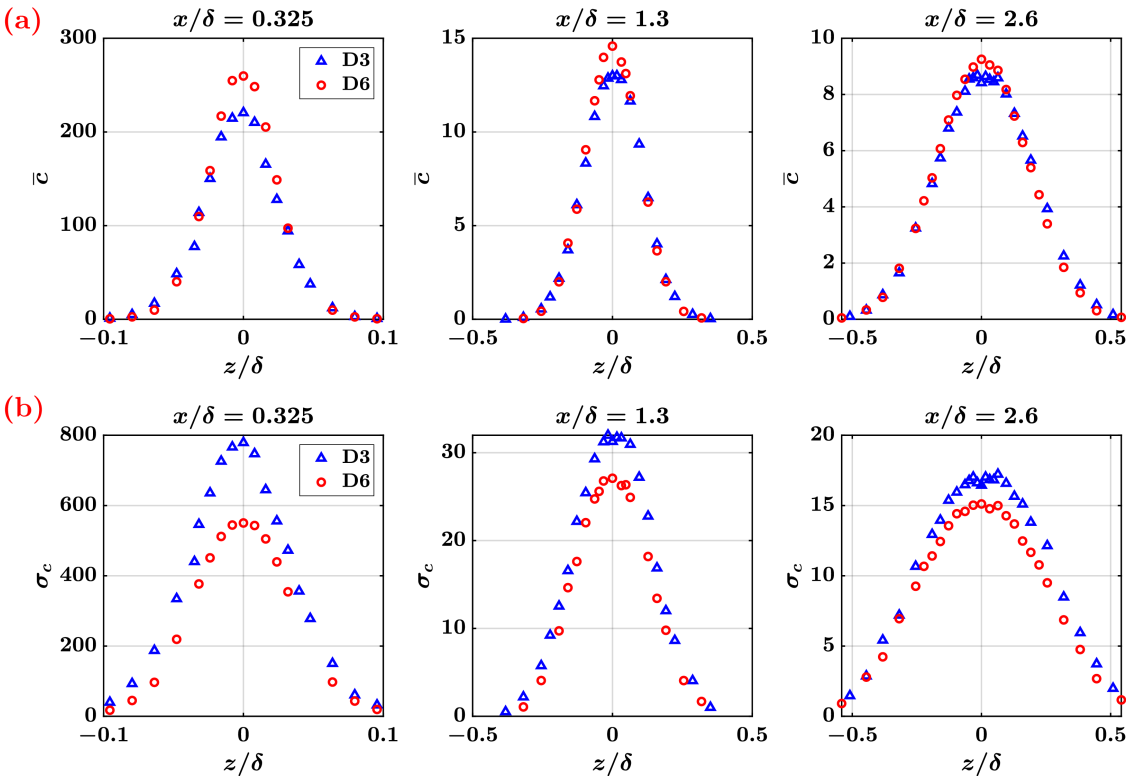


Figure 5.25: Transversal profiles of the mean and standard deviation values of the concentration, $\overline{c}$ and σ_c, at different streamwise locations, for the two source diameters, $D = 3$ mm and $D = 6$ mm.

Transversal profiles of passive scalar concentration

Additional results are reported for the time-series analysis of the passive scalar concentration. Differently from the main analysis (carried out in the previous Section 5.4.2) that focuses on the wall-normal direction, the first two moments of c as well as the corresponding network metrics, are here described as a function of the spanwise coordinate, for the two source diameters, D3 and D6.

The transversal profiles of mean and standard deviation of concentration, $\overline{c}$ and σ_c, are illustrated in Figure 5.25(a),(b) as a function of z/δ. We recall that the meandering motion acts both the in wall-normal, y, and spanwise, z, directions. Therefore, the mean concentration field (see Figure 5.25(a) and Figure 5.13(a)) is weakly affected by the source size, except in the proximity of the source. Conversely, the standard deviation of c is deeply affected by a different source size, either in the wall-normal (Figure 5.13(b)) and in the spanwise (Figure 5.25(b)) directions. It can be observed that even at large distance from the source (e.g., see the right panel in Figure 5.25(b)) σ_c for D3 is notably higher than σ_c for D6.

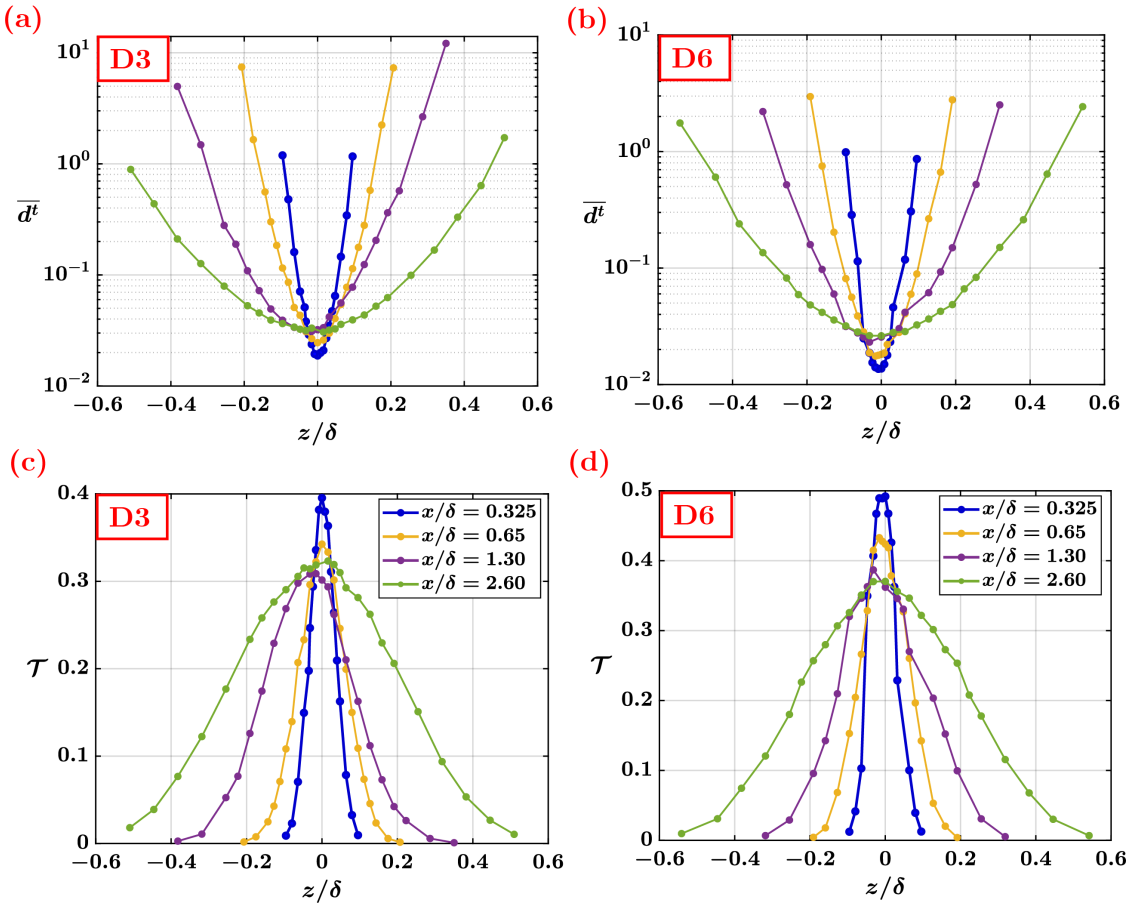


Figure 5.26: Network metrics as a function of the spanwise coordinate z/δ, for different streamwise locations, x/δ, and for D3 (left panels) and D6 (right panels). (a)-(b) Mean link-length, (c)-(d) transitivity.

From the point of view of visibility networks, Figure 5.26 shows the transversal behaviour of the (average) mean link-length, $\overline{d^t}$ (Figure 5.26(a)-(b)), and the transitivity, $\mathcal{T}$ (Figure 5.26(c)-(d)). Similarly to the analysis in the (x-y) plane (see Figure 5.18), at the source axis (i.e., $z/\delta = 0$) $\overline{d^t}$ and $\mathcal{T}$ show their minimum and maximum values, respectively. Moreover, the network metrics deeply change by moving away from the source axis, namely by increasing/decreasing z/δ. As for the wall-normal analysis, $\overline{d^t}$ for the smallest source diameter, D3, is generally larger than for D6, while $\mathcal{T}$ is generally smaller for D3 than for D6. As discussed in Section 5.4.2, the plume emitted by a smaller source is affected by a stronger meandering motion, which reflects in more intermittent signals of concentration (producing larger $\overline{d^t}$ and small $\mathcal{T}$). However, it is worth highlighting that, differently from the analysis in the (x-y) plane, maximum and minimum of the metrics are precisely found at the source axis ($z/\delta = 0$) as there is no effect of the mean

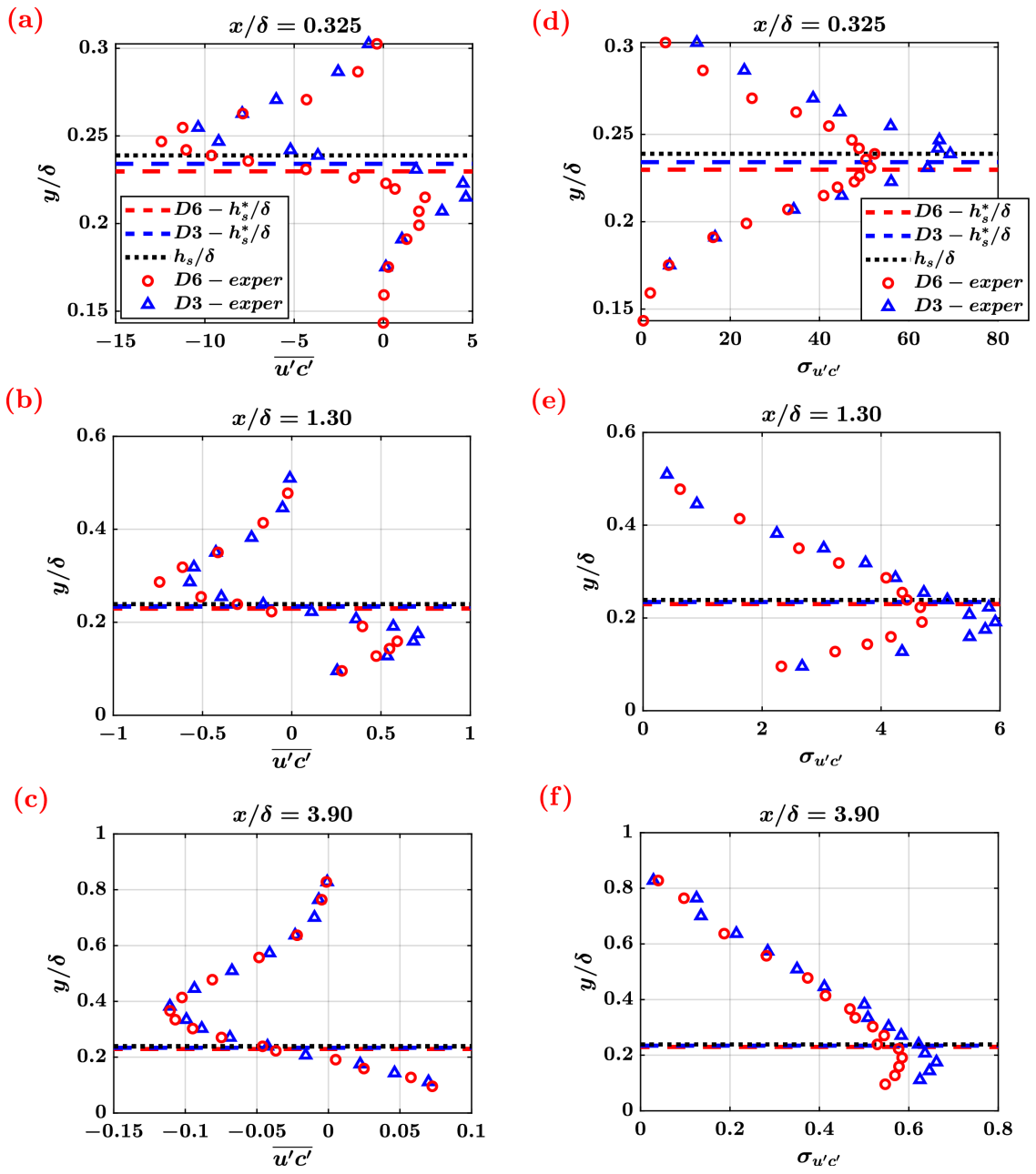


Figure 5.27: Vertical profiles of the mean value of longitudinal transport $\overline{u'c'}$ (a)-(c), and the standard deviation $\sigma_{u'c'}$ (d)-(f). The profiles are plotted at $x/\delta = \{0.325, 1.30, 3.90\}$, for the two source diameters, D3 and D6. The source axis height, h_s, is illustrated as a horizontal dotted line, while the plume axis height, h_s^*, is displayed as a blue (red) dashed line for the source D3 (D6).

shear in the transversal (x-z) plane. Furthermore, the behaviour or the metrics shown in Figure 5.26 is symmetrical with respect to $z/\delta = 0$, since there is not

a boundary effect due to the presence of the wall in the spanwise direction (for comparison, see the metric asymmetry in Figure 5.18).

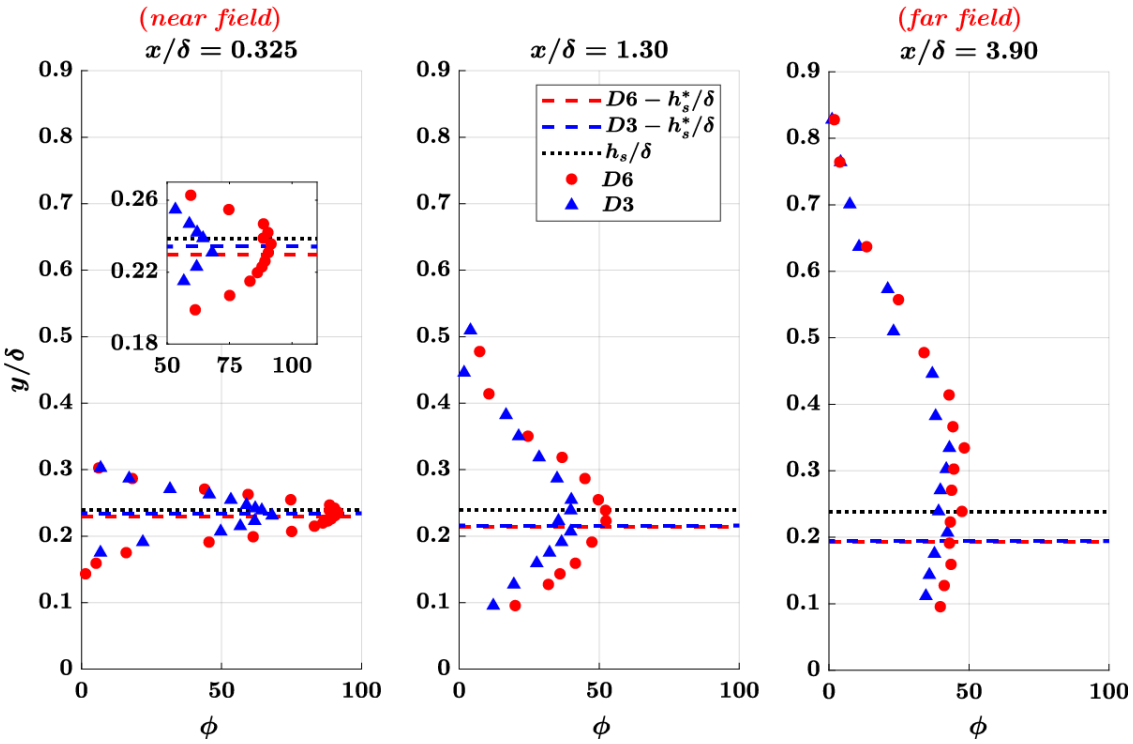


Figure 5.28: Vertical profiles of the average peak occurrence ϕ [Hz] based on signals of $u'c'$. The profiles are plotted in the near field and in the far field, for D3 (blue marker) and D6 (red marker). The source axis height, h_s, is illustrated as a horizontal dotted line, while the plume axis height, h_s^*, is displayed as a blue (red) dashed line for the source D3 (D6). The inset at $x/\delta = 0.325$ shows a zoom around the plume axis.

Vertical profiles of longitudinal turbulent transport

Figure 5.27 shows the vertical profiles of the mean value, $\overline{u'c'}$ and the standard deviation, $\sigma_{u'c'}$, of the (normalized) longitudinal turbulent flux, $u'c'$. The profiles are reported at three representative streamwise locations, i.e. in the near field ($x/\delta = 0.325$), in the far field ($x/\delta = 3.90$), and at an intermediate location ($x/\delta = 1.30$). As for the concentration field and vertical turbulent transport, the mean values of $u'c'$ are similar for the two source configurations D3 and D6 (see Figure 5.27(a)-(c)). It should be emphasized the difference in the vertical profiles of $\overline{u'c'}$ with respect to the vertical profiles of $\overline{v'c'}$ (see Figure 5.14(a)-(c)). In particular, above the source axis, positive v' fluctuations tend to carry the passive-scalar upwards, resulting in a positive $\overline{v'c'}$. On the other hand, negative v' fluctuations

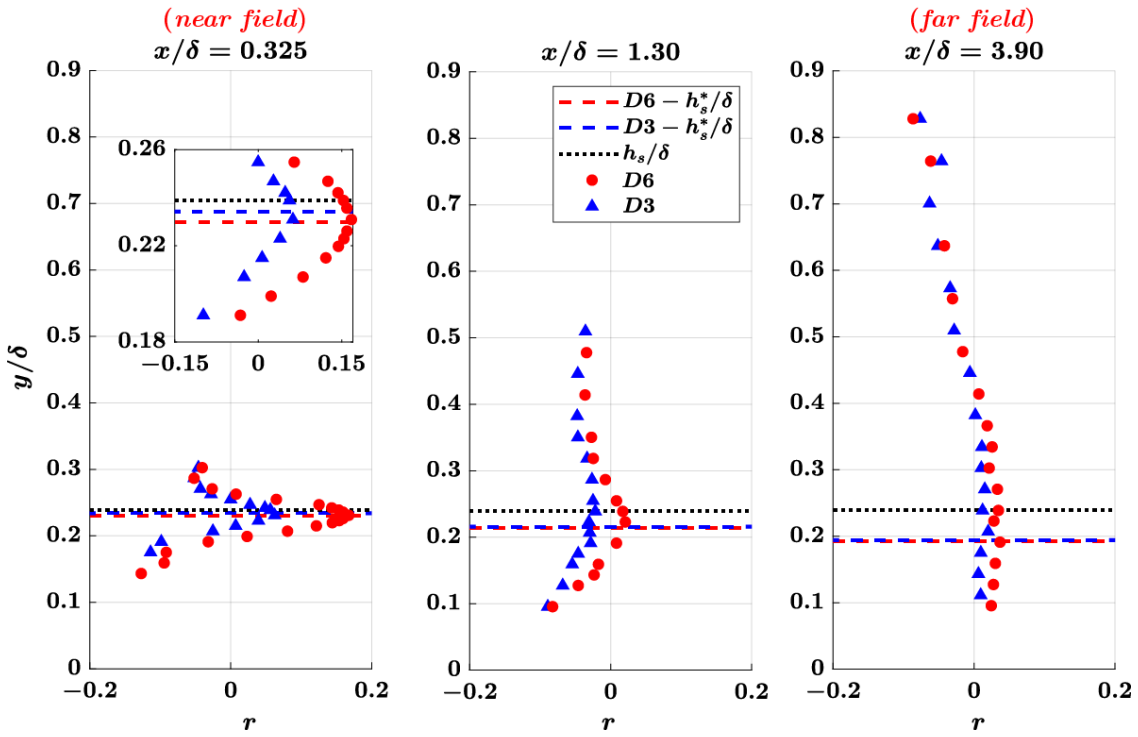


Figure 5.29: Vertical profiles of the assortativity coefficient r based on signals of $u'c'$. The profiles are plotted in the near field and in the far field, for D3 (blue marker) and D6 (red marker). The source axis height, h_s, is illustrated as a horizontal dotted line, while the plume axis height, h_s^*, is displayed as a blue (red) dashed line for the source D3 (D6). The inset at $x/\delta = 0.325$ shows a zoom around the plume axis.

tend to carry the passive scalar downwards. Since $\overline{u'v'} < 0$ in a TBL (see Figure C.2(c) in Appendix C), above the source axis $\overline{u'c'} < 0$, while below the source axis $\overline{u'c'} > 0$. The effect of the source size for an elevated source is much more evident for the standard deviation, $\sigma_{u'c'}$, even at large distances from the source (see Figure 5.14(d)-(f)). This is a consequence of the stronger meandering motion of the plume emitted by the smallest source size, D3, which produces more variability in the series. It is worth to observe that the overall values of $\sigma_{u'c'}$ are higher than the values of $\sigma_{v'c'}$, thus implying a stronger effect of the plume meandering in the signals of longitudinal turbulent transport (which are then expected to be more intermittent).

The application of the natural visibility algorithm to time-series of longitudinal turbulent transport is briefly discussed (similarly to the vertical transport) in terms of average peak occurrence, ϕ, and assortativity coefficient, r. As shown in Figure 5.28 and Figure 5.29, the average peak occurrence and the assortativity

coefficient extracted from $u'c'$ show several overall analogies with respect to the corresponding metrics for $v'c'$ (see Figure 5.20 and Figure 5.21). For instance, for both $u'c'$ and $v'c'$, the maximum values of the metrics are found in correspondence of the plume axis, h_s^*, and the metrics for the largest source, D6, are generally higher than the metrics for the smallest source, D3. As discussed previously, the difference in the profiles of the metrics for D3 and D6 is due to a different meandering intensity in the two source configurations.

However, there are some differences in the intensity and the trend of the two metrics whether $v'c'$ or $u'c'$ is examined. The average peak occurrence, ϕ, is always smaller for $u'c'$ than for $v'c'$, and this difference is more notable in the near field than in the far field. The smaller values of ϕ for $u'c'$ indicates that peaks are less frequent than for $v'c'$. This outcome is in accordance with the fact that, as mentioned above for the standard deviations, the time-series of longitudinal turbulent transport are more intermittent than the time-series of vertical turbulent transport. Concerning the assortativity coefficient, the magnitude of r is similar for both $u'c'$ than for $v'c'$ (see Figure 5.21 and Figure 5.29). However, a closer look at the profiles of r reveals that the assortativity coefficient is rather higher for the vertical turbulent transport, $v'c'$, than for $u'c'$, implying that outliers are more probable in the signals of $u'c'$. This is again in line with the fact that longitudinal turbulent transport signals are more intermittent than vertical turbulent transport ones.

5.5 Summary and future outlooks

The application of the visibility graph to time-series extracted from wall-bounded turbulent flows is tackled in this Chapter. First, novel insights are provided into how the network metrics are affected by the different temporal structure of the mapped time-series, in terms of irregularities and extreme events. The most suitable network metrics that are able to capture the specific features of temporal structure of time-series are the mean link-length (or average peak occurrence), the transitivity, the assortativity coefficient and the degree centrality. Two setups are considered, namely (i) the velocity field in a (numerically simulated) turbulent channel flow and (ii) the concentration and turbulent transport of a passive scalar plume in an experimental turbulent boundary layer.

In the turbulent channel flow, the presence of different topological features along the wall-normal direction is revealed by different metric trends. In particular, metrics variations are associated to different flow dynamics (responsible of the temporal structure of the signals) at various wall-normal coordinates. For the turbulent boundary layer setup, the analysis of concentration and turbulent transport of passive scalar reveals that the network metrics are significantly affected by measurement locations and the interplay between the meandering and the relative dispersion of the plume. Besides, a classical statistical analysis of vertical turbulent

transport of passive scalar is also provided. In this way, we extend the benchmark of (one-point) statistics of *Nironi et al.* [71] and *Fackrell and Robins* [77] on the dynamics of a passive scalar plume emitted in a rough-wall turbulent boundary layer.

Despite several statistical techniques are available to study nonlinear time-series, specifically regarding turbulence, most of them are invariant under different temporal structures of the time-series. The visibility-network analysis, instead, reveals to be a powerful and synthetic tool to handle *big-data* and to explore specific temporal features of the mapped series, without losing information about their temporal structures and also capturing the underlying flow dynamics. In fact, each network retains the temporal dislocation of important temporal features, such as extreme events and irregularities. Specifically, in terms of extreme events, the visibility network-based approach is able to discern between peaks and outliers and their frequency in the signals. To extract and handle this information is crucial for a deeper understanding of the flow dynamics, since the most common statistical tools adopted in turbulence, from spectral analysis to higher-order moments, are not able to capture the temporal collocation of such phenomena. Our visibility-based approach demonstrates that complex networks are able to provide important information about temporal structure of turbulent time-series, either in a numerical and an experimental setup.

Based on present findings, the proposed procedure can provide the basis for a novel way to approach time-series of turbulent flows, thus supporting classical statistical methods in time-series analysis. Future works can involve other turbulent flow configurations as well as data extracted at different Reynolds numbers, thus investigating the effects of the flow setup and Reynolds number on the scaling behaviour of the network metrics. Alternative turbulent quantities (e.g., energy or vorticity fields) can also be employed for mapping time-series in visibility networks, to highlight further aspects of turbulence dynamics that could not be easily unravelled by only focusing on the analysis of velocity components.

Chapter 6

Turbulent channel flows in the view of spatial networks

Some of the contents presented in this Chapter have been previously published. Below are provided the references:

S. Scarsoglio, G. Iacobello, and L. Ridolfi. "Complex networks unveiling spatial patterns in turbulence". In: *International Journal of Bifurcation and Chaos* 26.13 (2016), p. 1650223. DOI: 10.1142/S0218127416502230.

G. Iacobello, S. Scarsoglio, J.G.M. Kuerten, and L. Ridolfi. "Spatial characterization of turbulent channel flow via complex networks". In: *Physical Review E* 98.1 (2018), p. 013107. DOI: 10.1103/PhysRevE.98.013107. Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevE.98.013107

6.1 Motivation

The Eulerian approach has historically played a fundamental role for studying fluid flows, either experimentally or numerically. The idea to measure physical quantities at fixed spatial locations to extract insights into the flow dynamics, has paved the way to several fundamental results about the spatial characterization of turbulence. However, as remarked in Chapter 2, several issues regarding turbulence (especially if wall-bounded) still remain open.

Among the techniques employed, the correlation coefficient has been extensively used in the literature to characterize the spatio-temporal dynamics of turbulence. However, correlation coefficients are usually investigated by averaging their values in space and time, thus losing part of the spatial information comprised in each correlation matrix. Therefore, in order to advance the level of information of classical statistics, a spatial network-based approach to study turbulence from an Eulerian viewpoint is proposed. Although spatial networks based on correlation have been

largely studied in the context of climate networks, the application of network analysis to investigate 3D spatio-temporal turbulent data was not pursued before the present study. It is worth to note that, in order to build a spatial network, simultaneous measurements of physical quantities (e.g., the velocity components) are required at different spatial locations.

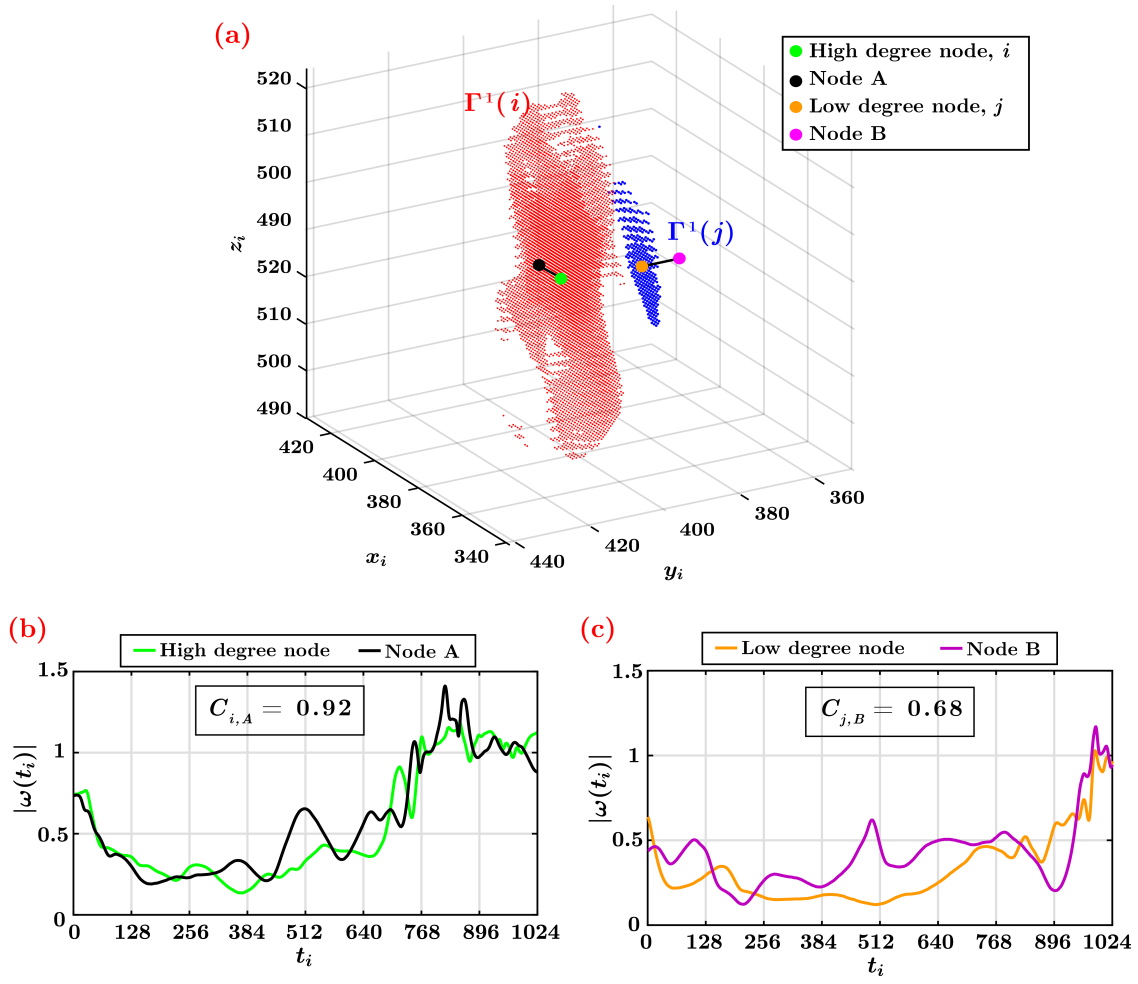


Figure 6.1: (a) High degree centrality node, i, and low degree centrality node, j, shown together with their first neighbourhoods ($\Gamma^1(i)$ in red and $\Gamma^1(j)$ in blue). Nodes A and B are at a distance equal to ten grid cells from i and j, respectively. Time-series of the vorticity modulus $|\omega(t_i)|$ are both shown for (b) the pair $\{i, A\}$ and (c) for the pair $\{j, B\}$ with the corresponding correlation coefficient, $C_{i,A} = 0.92$ and $C_{j,B} = 0.68$.

A homogeneous isotropic turbulent flow is firstly investigated by highlighting high-correlation values in the flow. The main idea is to assign nodes to grid points in

a spherical region of the total computational domain, and activate links by retaining the highest correlation between turbulent kinetic energy signals (see Ref. [148] for more details). Despite of the homogeneity and isotropy of the flow, we found that highly connected nodes (with threshold value equal to 0.9) tend to cluster in space, evidencing spatial patterns coherently moving with similar vorticity over one integral timescale. To better grasp the concept, Figure 6.1(a) shows a 3D view of the first neighbourhoods, Γ^1, of a node with high degree centrality, i (green dot), and another node with low degree centrality, j (orange dot); the neighbourhoods of i and j are illustrated as red and blue dots, respectively. Other two nodes labelled as A and B were focused and the vorticity time-series, $|\omega(t_i)|$, were extracted at the spatial locations corresponding to node pairs $\{i, A\}$ and $\{j, B\}$. Nodes A and B were selected to be at a distance equal to ten grid cells from i and j, respectively (note that A is a neighbour of i). The pair involving the high degree node presented a strong temporal correlation for the vorticity (see Figure 6.1(b)), while a weaker correlation was found for the node pair involving the low degree node (see Figure 6.1(c)). The behaviour of these node pairs is representative of high degree centrality and low degree centrality regions, since analogous comparisons were found for many other couples of nodes. Thus, we concluded that high degree centrality values indicate regions with the same instantaneous vorticity, that is, turbulent patterns coherently moving over the integral timescale. Therefore, although the network was built on the turbulent kinetic energy, the network is able to capture the spatial information on a higher-order level (i.e., the velocity gradients).

Motivated by the findings reported in Ref. [148], a spatial network-based analysis of fully-developed turbulent channel flow is pursued to offer an innovative approach to wall-turbulence. The results provided in the following sections of this Chapter, therefore, represent a first effort to characterize wall-bounded turbulent flows via complex network tools.

The Chapter is organized as follows. Section 6.2 reports the details of the spatial network construction. The main analysis of the spatial network built on the streamwise velocity component is discussed in Section 6.3. In particular, three points of view are adopted, namely a global scale (Section 6.3.1), a mesoscale (Section 6.3.2) and a local scale analysis (Section 6.3.3). Complementary results are reported in Section 6.4, which includes findings (i) obtained from a different Reynolds number and for different temporal windows (see Section 6.4.1), and (ii) from the network built on the wall-normal velocity (see Section 6.4.2). Finally, a resume of the results and future perspectives are illustrated in Section 6.5.

6.2 Building the turbulent spatial network

Velocity data from a fully-developed turbulent channel flow – numerically solved at a frictional Reynolds number $Re_\tau = 180$ (full details of the DNS are reported in Appendix B.3) – are exploited to construct a correlation-based spatial network. The streamwise velocity component, u, is the principal variable focused to carry out the network analysis of the channel flow. The u component, indeed, is one of the most significant variables to characterize a turbulent channel flow (see Section 2.2). Besides u, the wall-normal velocity component, v, is also investigated for comparison in Section 6.4.2. Therefore, in the following, unless otherwise stated, the procedure and results are intended for the streamwise velocity component.

An example of the procedure carried out to map the spatio-temporal velocity field in a complex network is illustrated in Figure 6.2: the discretization of the domain and the velocity field are illustrated in Figure 6.2(a), while the resulting correlation-filtered spatial network is shown in Figure 6.2(b). In particular, the spatial network is built – as discussed in Section 4.3.2 – by following four steps:

1. a network node is assigned to each grid point in the domain;
2. time-series of the velocity at each grid point are employed to evaluate the (Pearson) correlation coefficient, $C_{i,j}$, between all pair of nodes;
3. a threshold operation is performed by retaining all the correlation coefficients (in modulus) greater than a specific threshold value, namely $|C_{i,j}| > \theta$;
4. a link between two nodes is established if the correlation coefficient between velocity signals satisfies the previous threshold criterion.

In what follows, each step is discussed in more detail.

Node assignment. The velocity time-series are extracted in $N_x = 144$, $N_y = 191$ and $N_z = 150$ grid points along the streamwise, wall-normal and spanwise directions, respectively. By assigning a node to each selected grid point, $\mathcal{N}_v = N_x \times N_y \times N_z = 4125600$ nodes are obtained. It should be noted that the size and spatial discretization of the domain in the homogeneous directions are reduced in order to have a more manageable network size, without substantially affecting the statistical significance of the results (see Section B.3 for more details). Since the grid points are not uniformly distributed along the wall-normal direction, nodes represent an heterogeneous set of spatial entities. Therefore, a weight indicating its spatial extension is assigned to each node, namely the volume, $\mathcal{V}_i$, represented by each grid point (e.g., see Figure 6.2(a)). As a result, nodes at different y^+ coordinates have different weights equal to $\mathcal{V}_i(y^+) = (\Delta x \cdot \Delta y_i(y^+) \cdot \Delta z)$. While in the streamwise and spanwise directions the spacings are uniform (i.e., Δx and

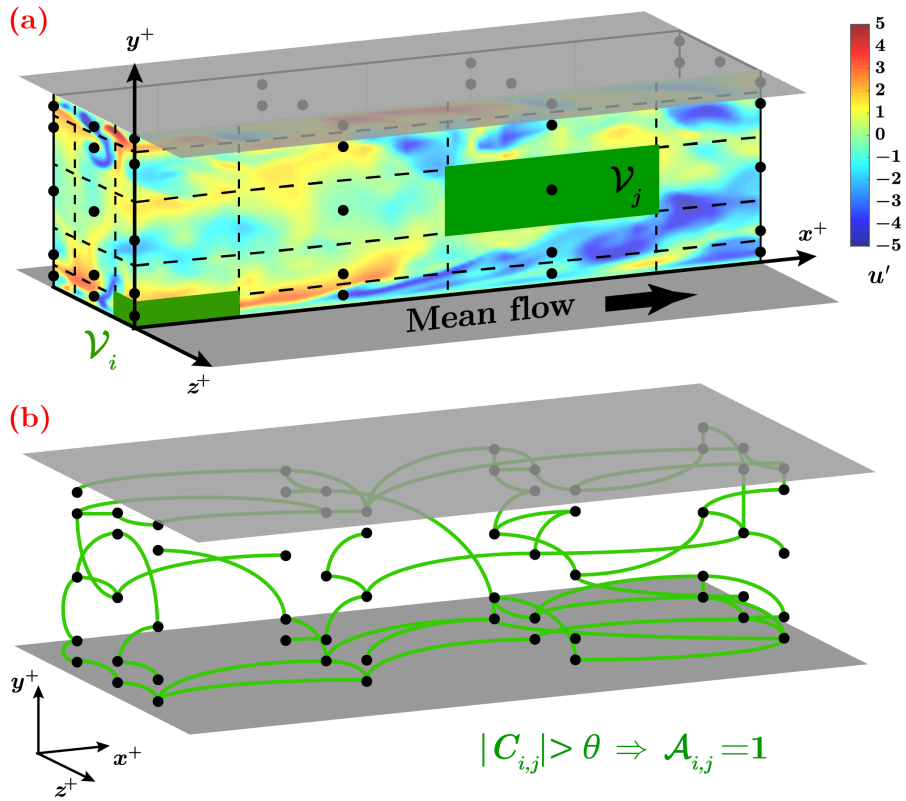


Figure 6.2: (a) 3D view of the turbulent channel setup. The streamwise velocity fluctuation, u', is shown in a $(x-y)$ and a $(z-y)$ section. An example of grid is also illustrated as black dots, by highlighting two cell volumes, $\mathcal{V}_i$ and $\mathcal{V}_j$, in green. (b) An example of 3D spatial network built from the setup in (a).

Δz are constant), in the wall-normal direction the spacing depends on y^+ through the relation $\Delta y_i = (y_{i+1} - y_{i-1})/2$. In other words, $\Delta y_i(y^+)$ is calculated as the sum of the previous, $i-1$, and next, $i+1$, half-heights of the grid spacing in the y direction.

Correlation computation. The Pearson correlation coefficients, $C_{i,j}$, are computed for each pair of nodes and for a total simulation time $T^+ = 225$, which corresponds to about 1.5 times the flow through time (L_x^+/U_b^+, where U_b is the bulk velocity) and about 11 times the integral time-scale. Specifically, $C_{i,j}$ is evaluated by following the Eq. (4.3). In order to speed-up the computation, only half the correlation matrix is evaluated, due to the symmetry property of $C_{i,j}$. Figure 6.3 shows the probability density function (PDF) of the values of the correlation coefficient of u for the turbulent channel flow investigated. The PDF is symmetrical for small $|C_{i,j}|$ values, approximatively in the range $-0.7 < C_{i,j} < 0.7$. However, as shown in Figure 6.3(b), for high values of $|C_{i,j}|$ the fraction of positive correlation

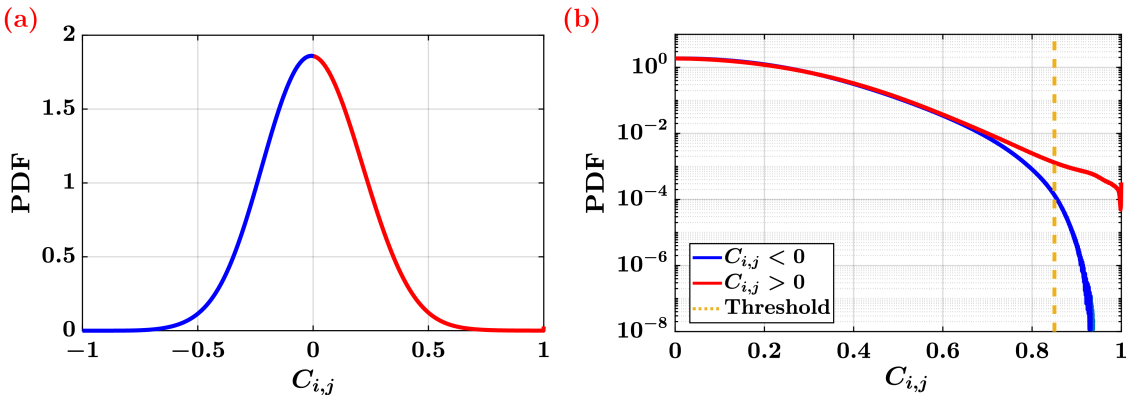


Figure 6.3: Probability density function (PDF) of the correlation coefficient values in a Linear-Linear plot (a) and a Linear-Log plot (b). Blue lines indicate negative correlations while red lines indicate positive values. In panel (b) a threshold line for $\theta = 0.85$ is also shown as a orange dashed line.

coefficients is larger than the fraction of negative correlations. It is remarkable to notice that, although rare, strong negative correlation values are present.

Correlation threshold. In order to activate links between nodes, the correlation coefficient is filtered by retaining its highest values. Constructing networks by filtering correlation coefficients is a classical technique employed in network analysis [158, 182]. In this work, both positive and negative high correlation values are considered by taking $|C_{i,j}| > \theta = 0.85$. The choice of the threshold is a non-trivial aspect for correlation-based networks. In fact, a θ value that is too high leads to extremely sparse networks (see Figure 6.3(b)), in which mainly trivial connections are unveiled. On the contrary, a too low value of θ results in networks where the statistical significance of the links is arguable (as there are *too many* links), thus making the interpretation of the network structure confused or misleading. Consequently, a reference value of the threshold is chosen as $\theta = 0.85$ to highlight the strongest (linear) relations between nodes at different spatial locations. In order to explore the sensitivity of the spatial network to the choice of θ, a parametric analysis of the results for different threshold values, $\theta = \{0.8, 0.9, 0.95\}$, is also reported throughout the results. Therefore, unless purposely stated, the analysis is intended for the reference case with $\theta = 0.85$.

Link activation. Links between node pairs – namely, between different spatial locations in the channel domain – are established if the correlation coefficient values satisfy the thresholding condition, $|C_{i,j}| > \theta$ (e.g., see Figure 6.2(b)). The choice of a high θ value allows for a total number of links, $\mathcal{N}_e$, that is computationally manageable. In particular, for $\theta = 0.85$ the number of links in the network

is $\mathcal{N}_e = 857693107 \sim 10^8$ so that the corresponding network edge density (see Eq. (3.4)) is $\varrho \approx 10^{-4}$, as the maximum number of possible edges in a network is $\mathcal{N}_v(\mathcal{N}_v - 1)/2 \sim 10^{12}$. The value of ϱ is very low, meaning that the network (as expected) is sparse, which is a common feature in several real-world networks.

To summarize, a spatial network based on the velocity time-series is built. Nodes represent physical portions of the domain and node weights are proportional to the cell volume of each grid point. Links indicate intense (positive or negative) correlations between different spatial locations, thus highlighting the spatial structure of the strongest kinematic similarities (i.e., in terms of velocity) between nodes.

6.3 Multi-scale network analysis of the streamwise velocity field

Results of the network analysis are presented in this section to highlight how the kinematic information spatially flows in the temporal window considered, and how this kinematic information is organized at three different scales:

(i) ***Global scale***. The overall characteristics of the whole network are investigated (i.e., considering all nodes, without any distinction); in particular, the network centrality and its assortativity behaviour are explored.

(ii) ***Mesoscale***. The attention is given to the topological features of groups of nodes; specifically, we study the network metrics as a function of two key features: the wall-normal coordinate, y^+, and the nodes with the highest centrality index (i.e., the most important ones).

(iii) ***Local scale***. The analysis is focused on single nodes and their relation with other nodes in the network via shortest paths.

The analysis at different scales allows to characterize the spatial network at different levels of understanding, from a macro-level that involves the entire network to the behaviour of single nodes. In this way, it is possible to fully exploit the potential of complex networks to capture information on the turbulent flow features.

6.3.1 Global scale analysis

The global behaviour of the spatial network is first studied by investigating the volume-weighted connectivity, $\mathcal{K}^{vw}$ (see Eq. (3.6)). We recall that the volume-weighted connectivity of a node, i, is bounded in $(0, 1]$ and represents the fraction of volume connected to i. In other words, $\mathcal{K}^{vw}(i)$ is a centrality metric that quantifies

the extent to which a spatial region of the domain (identified by a node i) is positively or negatively correlated with the remaining domain.

First, the probability to find a node with a given value of $\mathcal{K}^{vw}$ is focused. To this aim, the cumulative probability distribution of $\mathcal{K}^{vw}$, namely $P(\mathcal{K}^{vw})$, is shown in Figure 6.4(a) for the reference case with $\theta = 0.85$. Specifically, we plot $1-P(\mathcal{K}^{vw})$ in order to highlight the tail of the distribution, i.e. the nodes with highest centrality values. The probability to have higher values of $\mathcal{K}^{vw}$ decreases exponentially, suggesting there is a relatively small number of nodes that are strongly connected with respect to the other nodes, thus representing the *hubs* of the network. Therefore, $P(\mathcal{K}^{vw})$ can be exploited to classify the centrality of nodes in the network. We refer to nodes with high $\mathcal{K}^{vw}$ values as a $H-\mathcal{K}^{vw}$ nodes (where H stands for *high*) if they satisfy the condition $1-P(\mathcal{K}^{vw}) \leq 10^{-2}$ (corresponding to the 99th percentile). On the contrary, we refer to nodes with low $\mathcal{K}^{vw}$ values as $L-\mathcal{K}^{vw}$ nodes (where L stands for *low*), indicating that they satisfy the condition $1-P(\mathcal{K}^{vw}) = 99 \times 10^{-2}$ (corresponding to the 1st percentile). As a result, $H-\mathcal{K}^{vw}$ nodes represent parts of the domain kinematically similar to large portions of the physical domain, in the temporal window considered.

Figure 6.4(b) shows the effect of the threshold, θ, on the probability distribution of $\mathcal{K}^{vw}$. The cumulative $\mathcal{K}^{vw}$ distribution maintains a decreasing exponential behaviour for different values of θ, confirming the presence of hubs in the network. However, the slope of the probability distribution increases (in modulus) for increasing θ values. In fact, since very high correlations are unlikely to appear, the probability to find higher $\mathcal{K}^{vw}$ values tend to decrease for higher thresholds, θ.

$\mathcal{K}^{vw}$ is a measure of the centrality of nodes in the network, but it is not able to quantify whether the centrality of a node is similar or not to the centrality of its first neighbours. To address this issue, an assortativity metric (see Section 3.3.2) can be exploited based on a centrality metric, which is $\mathcal{K}^{vw}$ in this case. Assortativity metrics, in fact, are able to reveal if nodes tend to link to other nodes with similar or dissimilar $\mathcal{K}^{vw}$ values. Specifically, two assortativity metrics are here focused: the average nearest-neighbours connectivity (see Eq. (3.10)) and the assortativity coefficient (see Eq. (3.9)). The average nearest-neighbours connectivity, $\mathcal{K}^{vw}_{nn}(i)$, is the average value of $\mathcal{K}^{vw}$ of neighbours of a generic node i. If $\mathcal{K}^{vw}_{nn}(i)$ is an increasing (decreasing) function of $\mathcal{K}^{vw}(i)$ the network is assortative (disassortative). Figure 6.5(a) shows the average nearest-neighbours connectivity, $\mathcal{K}^{vw}_{nn}(i)$, as a function of $\mathcal{K}^{vw}(i)$, for each node i in the network. An almost linear relationship holds between $\mathcal{K}^{vw}$ and $\mathcal{K}^{vw}_{nn}$, displaying that most nodes tend to link to other nodes with quite the same $\mathcal{K}^{vw}$ value, resulting in a strongly assortative network. This behaviour is corroborated by the assortativity coefficient, which is equal to +0.84 (in Eq. (3.9) the volume-weighted connectivity is used instead of the degree). This outcome implies that parts of the domain with (linearly) similar time-series of the streamwise velocity u (i.e., high correlation coefficients values) have also similar neighbourhood spatial extensions. In other words, the fraction of

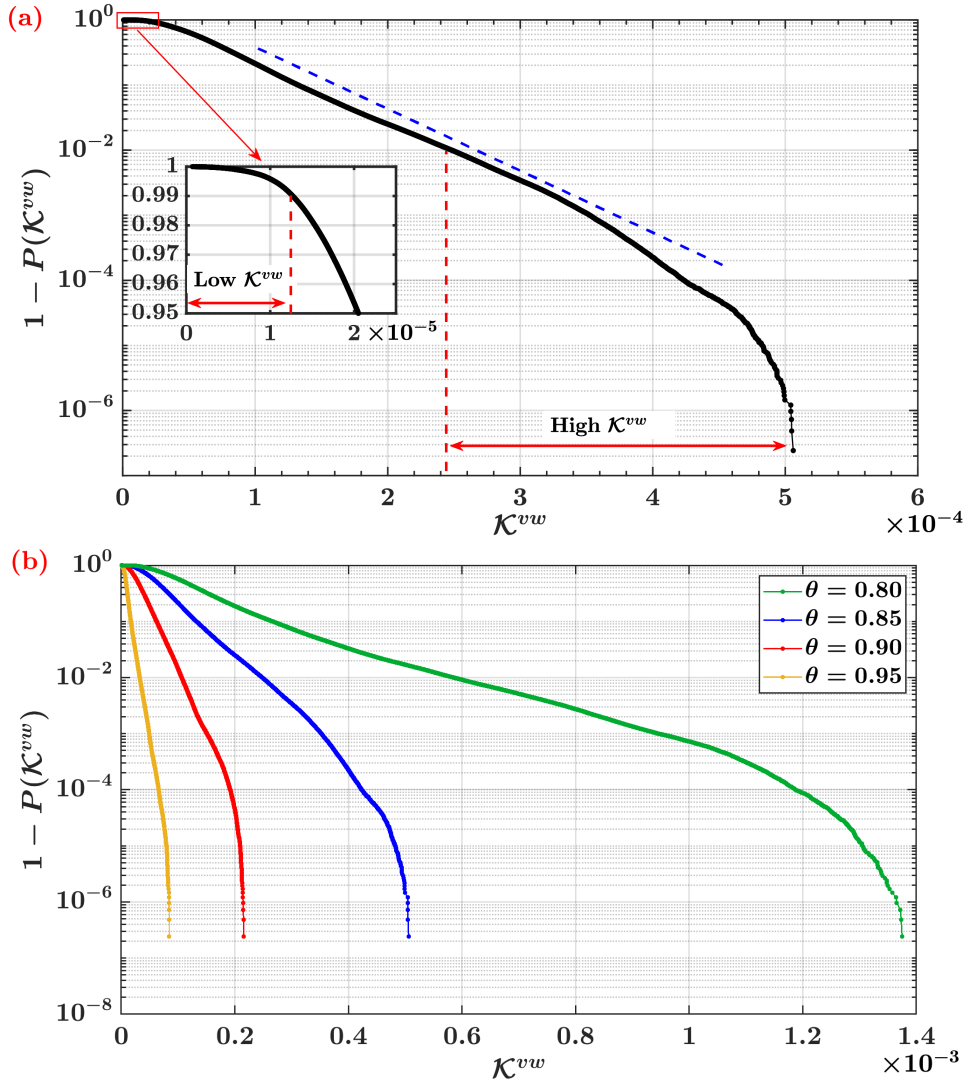


Figure 6.4: (a) Cumulative $\mathcal{K}^{vw}$ distribution, $1 - P(\mathcal{K}^{vw})$, and exponential trend for $\theta = 0.85$. The inset is a zoomed view for small $\mathcal{K}^{vw}$ values, indicated by the red box. The ranges of high-$\mathcal{K}^{vw}$ and low-$\mathcal{K}^{vw}$ (shown in the inset) are highlighted. (b) Cumulative $\mathcal{K}^{vw}$ distribution for different thresholds $\theta = \{0.80, 0.85, 0.90, 0.95\}$.

volume of the domain highly correlated with a node i and the fraction of volume highly correlated with each of the first neighbours of i (i.e., the nodes linked to i), are of the same order of magnitude.

The effect of the threshold value on the assortativity of the network is shown in Figure 6.5(b). An almost linear relation between $\mathcal{K}^{vw}_{nn}$ and $\mathcal{K}^{vw}$ holds for the networks at different θ, namely the correlation-based networks always display assortative behaviours. In particular, this effect is more evident for higher values of θ, as confirmed by the assortativity coefficient values that are equal to 0.78, 0.90 and 0.93

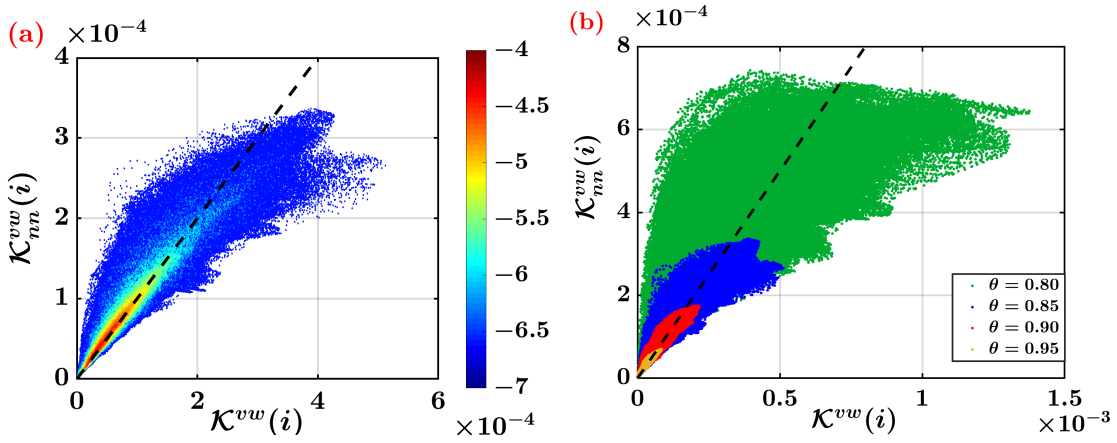


Figure 6.5: (a) Weighted average nearest neighbours assortativity measure, $\mathcal{K}^{vw}_{nn}(i)$, as a function of $\mathcal{K}^{vw}(i)$ for $\theta = 0.85$. Colours indicate the joint probability values (in $\log_{10}$ scale) of variables $\mathcal{K}^{vw}_{nn}(i)$ and $\mathcal{K}^{vw}(i)$. The bisector is also displayed as a black dashed line. (b) $\mathcal{K}^{vw}_{nn}(i)$ as a function of $\mathcal{K}^{vw}(i)$ for different threshold values $\theta = \{0.80, 0.85, 0.90, 0.95\}$.

for θ equal to 0.8, 0.9 and 0.95, respectively.

To summarize, the analysis at global scale points out that hubs are generally rare in the networks, especially for high threshold values. Additionally, nodes tend to connect with each others, namely the network is assortative for any threshold considered.

6.3.2 Mesoscale analysis

In order to deepen the features of the complex network structure, we move from a global to a mesoscale level of analysis. Namely, we focus on two specific sets of nodes according to two key features:

1. nodes are grouped according to their wall-normal coordinate, y^+, to highlight the network structure at different distances from the wall;

2. nodes are grouped according to their importance in the network in terms of volume-weighted connectivity, $\mathcal{K}^{vw}$, thus shedding light on the hierarchical structure of network nodes.

In the following part of this section, these two points of view are described in detail. In particular, due to the symmetrical behaviour of the results with respect to the center of the channel, the plots of the metrics as a function of y^+ are shown as averages of both halves of the channel (i.e., $y^+ \in [0, 180]$ and $y^+ \in [180, 360]$).

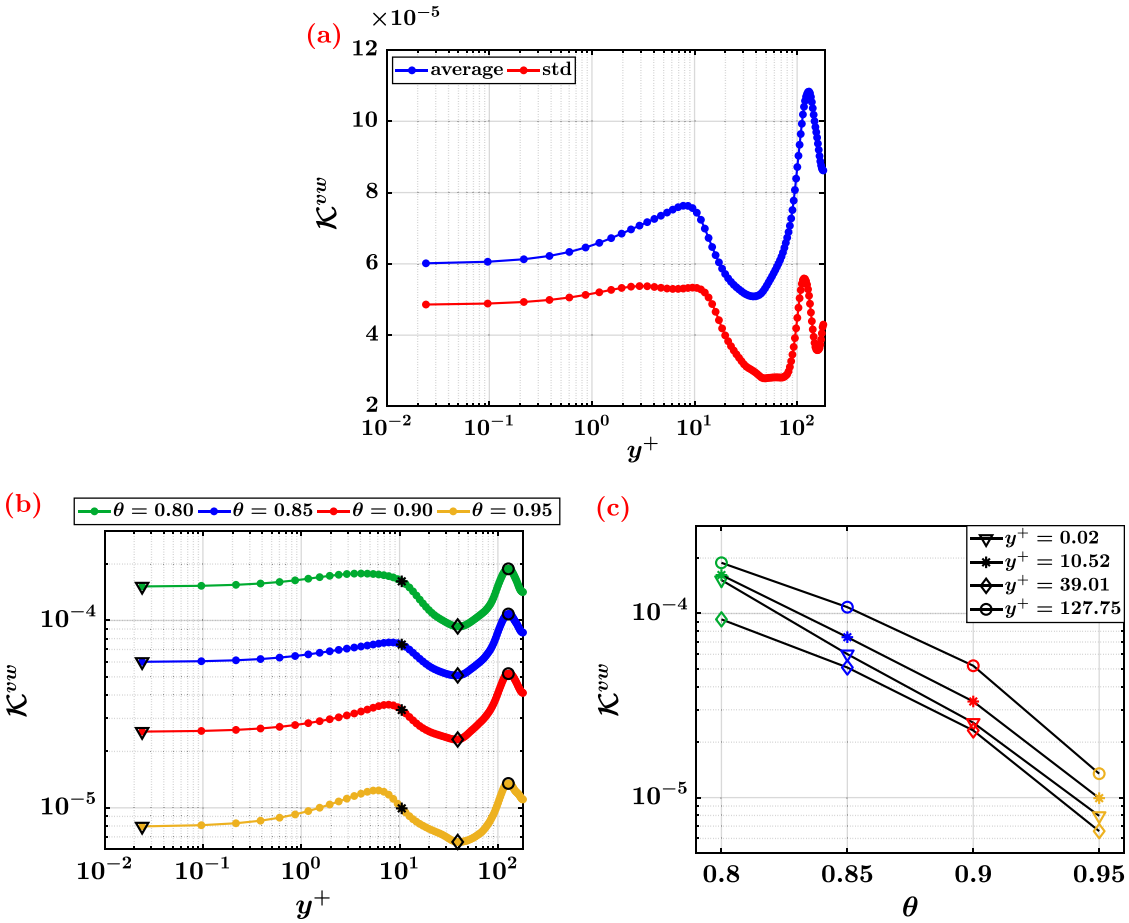


Figure 6.6: (a) Mean and standard deviation values (averaged over the two homogeneous directions) of the volume-weighted connectivity, $\mathcal{K}^{vw}$, as a function of y^+ for $\theta = 0.85$. (b) Mean values (averaged over the two homogeneous directions) of $\mathcal{K}^{vw}$ as a function of y^+ for different threshold values $\theta = \{0.80, 0.85, 0.90, 0.95\}$. (c) $\mathcal{K}^{vw}$ trends at fixed y^+ coordinates (see panel (b)), as a function of θ.

Vertical network structure

First, the mean and standard deviation values of $\mathcal{K}^{vw}$ evaluated in planes at constant y^+ are shown in Figure 6.6(a). We found two local maxima of the mean $\mathcal{K}^{vw}$ at distances from the wall of about $y^+ \approx 10$ and at about $y^+ \approx 120$. Such local peaks of the average value suggest the presence of a large number of $H - \mathcal{K}^{vw}$ nodes around those locations, namely nodes highly connected with the remaining network. Moreover, local maxima of the standard deviation are found at almost the same y^+ as the local peaks of the average value. This implies that $H - \mathcal{K}^{vw}$ nodes increase the variability of the $\mathcal{K}^{vw}$ values at these wall-normal locations. Therefore, hubs are rare in the network (as shown in Figure 6.4), and they are not evenly distributed in

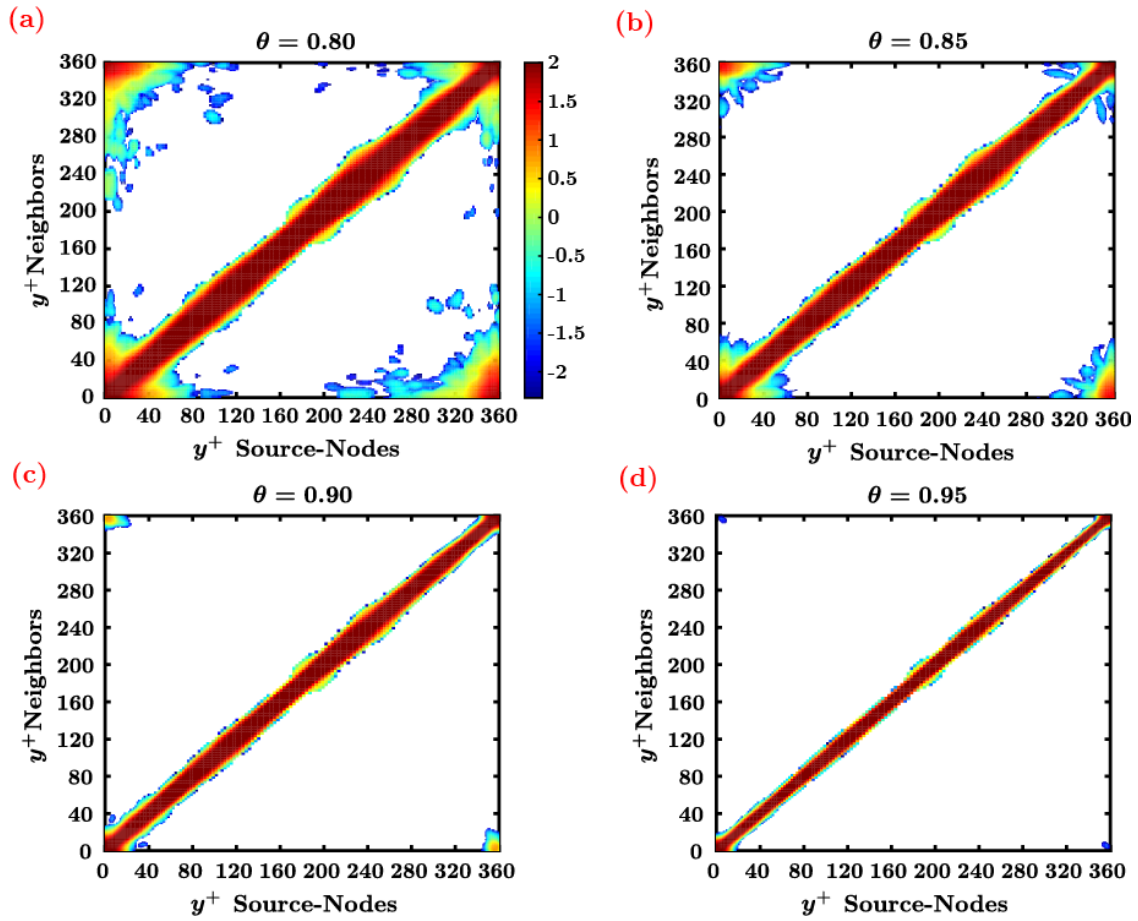


Figure 6.7: The probability (in percentage and in $\log_{10}$ scale) that a source-node at a given y^+ is linked to a neighbour at another y^+ value, for different threshold values θ. The color-scale ranges from $\log_{10}(100 \times (N_x N_z)^{-1}) \approx -2.34$ to $\log_{10}(100) = 2$ (i.e., 100% probability), where $(N_x N_z) = 21600$ is the number of nodes in a plane at fixed y^+.

the domain but tend to *cluster* at specific spatial locations. Figure 6.6(b) illustrates the effect of the correlation threshold, θ, on the mean $\mathcal{K}^{vw}$ profile. As θ increases, the average value of $\mathcal{K}^{vw}$ decreases at all wall-normal locations, as it is less probable to find high correlation values for large θ (namely, the nodes are less connected on average). In particular, mean $\mathcal{K}^{vw}$ values scale in a similar way for different y^+ by following an exponential decreasing behaviour as a function of θ. This is shown in Figure 6.6(c), which displays the mean $\mathcal{K}^{vw}$ values for four representative y^+. It is remarkable to notice that the location of local peaks of the mean $\mathcal{K}^{vw}$ profiles remains almost unchanged for different θ values.

Next, the probability that an arbitrary *source-node* at a fixed y^+ plane has a neighbour (i.e., is directly linked to another node) at others y^+ values is investigated.

In other words, we ask which is the fraction of connections between nodes located at two wall-normal coordinates. Figure 6.7 shows such a probability (illustrated in percentage and in a $\log_{10}$ scale) for each threshold value, θ, considered. Source-nodes at any y^+ have first neighbours close to themselves (diagonal part of the plot), for any θ value. However, we find that nodes corresponding to grid points close to one wall have non-zero probability to connect with other nodes that are also near the other wall. This outcome is rather evident for lower threshold values (i.e., $\theta = 0.8, 0.85$) while it tends to vanish for higher threshold values (i.e., $\theta = 0.9, 0.95$). Furthermore, by increasing the correlation threshold, nodes at any y^+ tend to have their first neighbours closer and closer to themselves in the wall-normal direction (i.e., the diagonal part of the plot stretches), but intra- and inter-wall links are still present for $\theta = 0.95$ for source-nodes very close to the wall (see Figure 6.7(d)). In particular, for the reference case at $\theta = 0.85$ (see Figure 6.7(b)), only nodes at a distance $y^+ \lesssim 70$ from one wall have non-zero probability values also near the other wall. Therefore, we can conclude that in the network built on the u component there are wall-wall links (both between nodes close to the same wall and at different walls) and center-center links, but there are no direct wall-center connections.

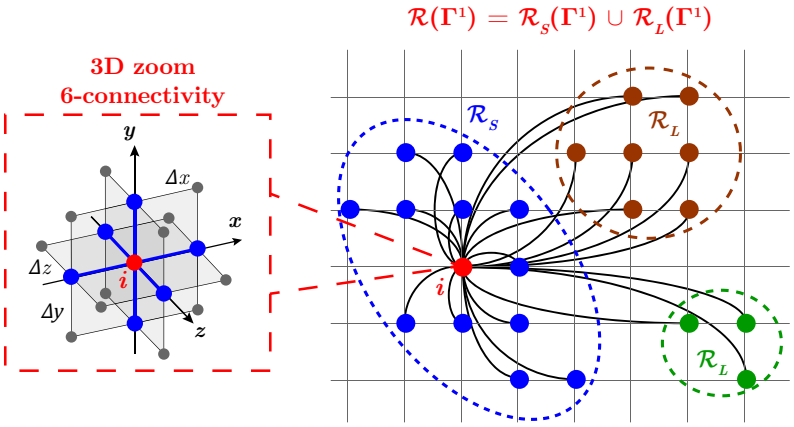


Figure 6.8: Sketch of the node i (red dot) and its first neighbourhood, $\Gamma^1(i)$, which is partitioned in a short-range region, $\mathcal{R}_S$ (blue dots), and two long-range regions, $\mathcal{R}_L$ (brown and green dots). A 3D zoom of the grid in the proximity of node i is also illustrated, to highlight the 6-orthogonal connectivity in the Cartesian discretization for the definition of region.

To further shed light on the appearance of inter-wall connections, we investigate the relation between a node, i, and its first neighbours, $\Gamma^1(i)$, namely the set of nodes directly linked to i. The spatial position of the first neighbours of a node i is studied, with the aim to explore if first neighbours spread all over the domain or there is some kind of spatial organization. With this aim, nodes are grouped according to their spatial locations. In fact, since to each node of the network corresponds a volume in a fixed spatial grid position, nodes that are close in space

can be grouped according to a connectivity criterion. More in detail, we say that a given set of nodes forms a *spatially-connected region* (or simply a *region*) if each node in the set is distant one grid spacing (in any Cartesian direction, $\pm\Delta x$, $\pm\Delta y(y^+)$ or $\pm\Delta z$) from at least another node of the set [42]. In the specific case in which the set of nodes to be grouped corresponds to the n-th neighbourhood, $\Gamma^n(i)$, of a node i, then we indicate each region of Γ^n as $\mathcal{R}(\Gamma^n)$. For example, Figure 6.8 shows a sketch of the six orthogonal connectivity criterion and the grouping in regions of the first neighbourhood of a node, i. In particular, if the first neighbourhood, $\Gamma^1(i)$, a generic node i is focused, we can distinguish between *short-range region*, $\mathcal{R}_S(\Gamma^1)$, and *long-range region*, $\mathcal{R}_L(\Gamma^1)$, of i. As a consequence, we say that a node j is a *short-range neighbour* of i if both i and j belong to the same spatially connected region, $\mathcal{R}_S(\Gamma^1)$ (see blue nodes grouped in the example of Figure 6.8). On the contrary, j is called a *long-range neighbour* of i, if j and i do not belong to the same region (e.g., see nodes in brown and green in Figure 6.8). By extension of terminology, we refer to short- and long-range links to indicate the connections between pairs of short- and long-range neighbours, respectively. By recalling that the number of elements in a set is here indicated via the $|\bullet|$ notation, for every node in the network it is true that $|\mathcal{R}_S(\Gamma^1)| = 1$ (since it can be only one region comprising a node) and $|\mathcal{R}(\Gamma^1)| = |\mathcal{R}_S(\Gamma^1)| + |\mathcal{R}_L(\Gamma^1)| = 1 + |\mathcal{R}_L(\Gamma^1)|$. The last relation points out that long-range neighbours are present only if $|\mathcal{R}(\Gamma^1)| > 1$, otherwise the first neighbourhood of a node i is only spatially connected to i (i.e., $\Gamma^1(i)$ is made up only by the short-range region). For instance, in Figure 6.8, the first neighbourhood of node i can be divided in one short-range region (blue nodes) and two long-range regions (green and brown nodes), i.e. $|\mathcal{R}(\Gamma^1)| = 2$. It should be noted that long-range neighbours of a node i are nodes *detached* from the short-range region, regardless of the physical (Euclidean) distance from i. Namely, there is a spatial gap (at least greater than one grid step, in each direction) that divides short- and long-range regions.

Figure 6.9(a) shows the mean and the standard deviation values of the number of regions, $|\mathcal{R}(\Gamma^1)|$, formed by first neighbours of nodes at fixed y^+. The first neighbourhood of nodes close to the wall tends to be composed of more than one region (as the mean is greater than one), while from $y^+ \approx 70$ up to the center of the channel, the first neighbours form only one region (with standard deviation equal to zero). This outcome is strictly in accordance with the results shown in Figure 6.7(b), which indicates that only nodes with coordinate $y^+ \lesssim 70$ display long-range connections. However, long-range regions are not only detected between nodes at different y^+ as shown Figure 6.7, but they are also present in the homogeneous directions, x and z. Figure 6.9(b) shows the mean and standard deviation of the weighted physical distance, $\overline{d^w}_\chi$ along the directions $\chi = x, z$, for nodes in planes at constant y^+. By following the definition given in Eq. (3.12), $\overline{d^w}_\chi$ is here evaluated only for nodes belonging to the long-range regions, $\mathcal{R}_L(\Gamma^1)$, of each node. In so doing, we focus on the average and variance of the distances

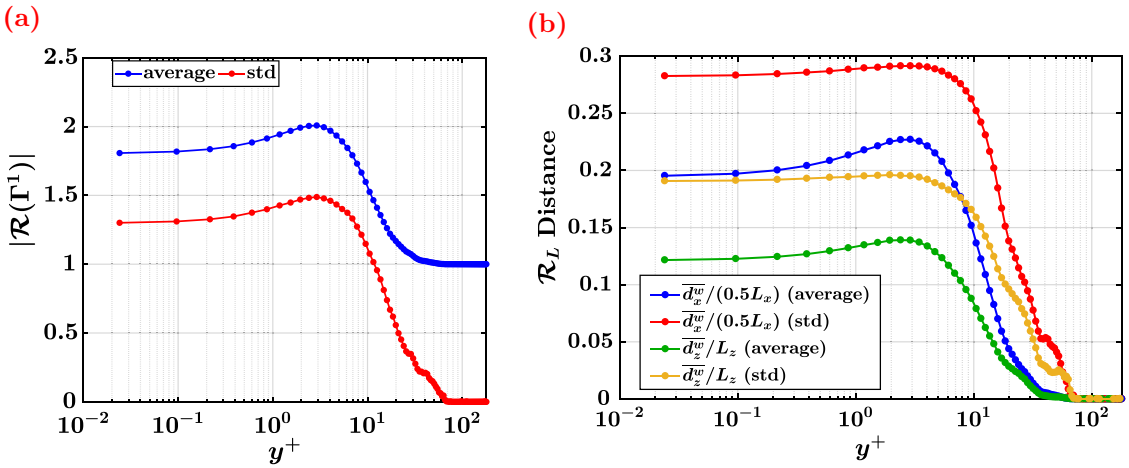


Figure 6.9: (a) Mean and standard deviation values of the number of regions, $|\mathcal{R}(\Gamma^1)|$, as a function of the distance to the wall, y^+, and averaged over the two homogeneous directions. (b) Mean and standard deviation of weighted physical distances between nodes at fixed y^+ and their long-range neighbours, averaged over the two homogeneous directions. The distances are normalized with the maximum distances, $0.5L_x = 2\pi$ and $L_z = 25/36\pi$, in the (x, z) directions. Due to the periodicity of the domain, the maximum distance in the x-direction is $L_x/2$ instead of L_x.

between a given node (located at a specific y^+) and its long-range neighbours. As shown in Figure 6.9(b), by moving from the center towards the wall, the long-range neighbours of a node tend to be located at increased distance (on average) in both the streamwise and spanwise directions. Moreover, long-range neighbours of nodes at the same y^+ are quite scattered in the x and z directions, as suggested by the high values of the standard deviation in Figure 6.9(b).

From the two-point spatial correlation definition [27, 44], it is straightforward expecting that some of the first neighbours of each node i in the network are located in the proximity of i, forming a short-range region, $\mathcal{R}_S$. This can be seen in Figure 6.7, where the highest probability values are in the diagonal part of the plots. Moreover, for $y^+ \gtrsim 70$, in Figure 6.9(a) the number of regions of the first neighbourhoods drops to one and in Figure 6.9(b) the distance between nodes and their long-range neighbours drops to zero. Instead, what is not trivial is the emergence of long-range links in all directions, more specifically inter- and intra-wall links occurring for $y^+ \lesssim 70$.

By analogy with the climate analyses, we refer to long-range links as *teleconnections* (see also Section 4.3.1). We recall that in atmospheric sciences teleconnections indicate climate relations (in terms of temperature, rainfall, pressure or other quantities) between geographically remote regions, farther than the correlation length

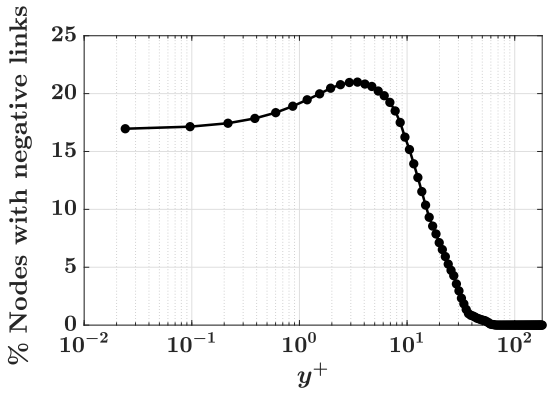


Figure 6.10: Fraction of nodes at constant y^+ with at least one negatively correlated link.

scale of the variable. In this work, the emergence of teleconnections of the streamwise velocity can be interpreted as the footprints of a top-down interaction, which similarly acts from the outer layer to both near-wall regions [45, 46]. In fact, teleconnections are always individuated between regions close to the two walls (or close to the same wall), revealing an analogous response of the two wall regions to the large-scale structures (i.e., turbulent structures with size of the order of the integral space scale). On the contrary, teleconnections are never found between inner and outer layer regions, where the interplay dynamics are deeply different one from each other. In other terms, here teleconnections do not emerge due to a direct relation between outer and inner layers, but due to an analogous response of the near-wall velocity field to a similar cause corresponding to large-scale motions. Therefore, complex networks are able to unveil the presence of teleconnections, which are usually hidden by the spatial averaging of the correlation coefficient values. Teleconnections create a texture of links (highlighted by the network metrics) between distant locations, in which similar (streamwise) kinematic information persists in time. This result is the main difference with respect to other approaches in the turbulence research, where the usual spatially-averaged correlation only retains average information about the spatial behaviour of the correlation field.

To complete the analysis as a function of y^+, the distribution of the sign of the correlation coefficient of links between nodes and their first neighbours is investigated. By hypothesis, in this work links in the network are active if the absolute value of the correlation coefficient, $C_{i,j}$, is above $\theta = 0.85$, but links can have either negative or positive $C_{i,j}$ values. To highlight the presence of links corresponding to negative correlations, the fraction of nodes at a given y^+ coordinate with at least a negative-correlated neighbour in the domain is shown in Figure 6.10. For instance, about 21% of the nodes at $y^+ = 3.5$ is highly negatively correlated with other nodes. Negative-correlation links are found only for $y^+ \lesssim 70$, with a peak

at $y^+ = 3.5$ that coincides with the peak of the average number of regions in Figure 6.9(a). More in detail, among the nodes with negatively correlated neighbours (i.e., nodes with non-zero percentages in Figure 6.10), the occurrence of negative links is (on average) about 10% of total links. This means that, by arbitrarily selecting a node i located at $y^+ \lesssim 70$, we expect that 1/10 of its first neighbours are negatively correlated with i. Therefore, based on what observed so far, we can infer that negative correlation links are possible due to the presence of teleconnections (i.e., $y^+ \lesssim 70$), while short-range links for $y^+ \gtrsim 70$ are only activated by positive correlation values.

Hierarchical network structure

The analysis of the network features as a function of the wall-normal coordinate, y^+, revealed two key points, namely (i) the most important nodes are not uniformly distributed throughout the domain, and (ii) nodes close to the walls display long-range links that are referred as teleconnections. The former result is focused in this Section, i.e. nodes are grouped according to their importance in the network in terms of volume-weighted connectivity, $\mathcal{K}^{vw}$. In particular, we recall that in Section 6.3.1 we referred as $H - \mathcal{K}^{vw}$ the nodes with $\mathcal{K}^{vw}$ value above the 99th percentile (i.e., the hubs of the network), while $L - \mathcal{K}^{vw}$ as nodes with $\mathcal{K}^{vw}$ below the 1st percentile.

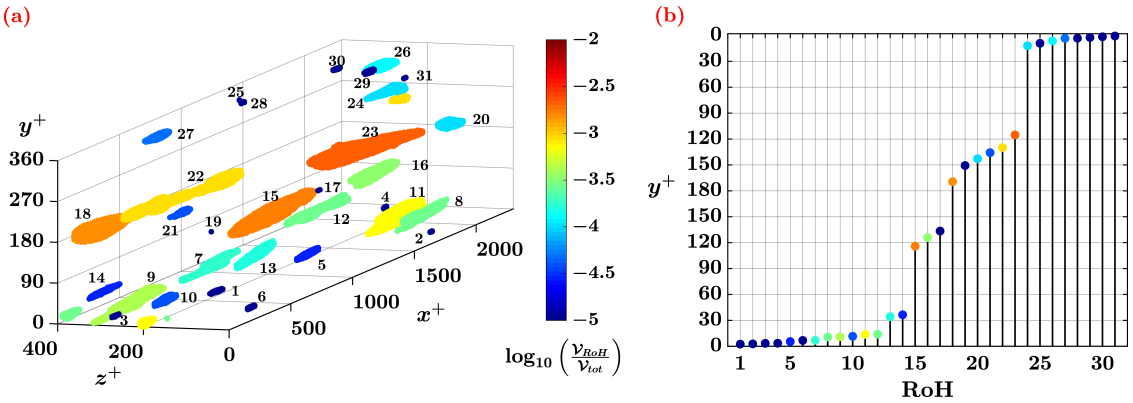


Figure 6.11: 3D view of high $\mathcal{K}^{vw}$ nodes, namely regions of hubs, RoHs, for $\theta = 0.85$. Colour scale refers to the fraction of volume occupied by distinct RoHs, $\mathcal{V}_{RoH}/\mathcal{V}_{tot}$. RoHs are labelled for increasing values of the wall-normal coordinate of their center of mass. Periodicity of the domain in the x direction is visible from the RoHs labelled 8, 11, 12 and 22. (b) Weighted y^+-component of the center of mass of the RoHs shown in panel (a), where the center of mass is evaluated as $y^+_{CoM} = \sum_i \left(y_i^+ V_i / V_{tot}\right) / \mathcal{V}(RoH)$, with $i \in$RoH and $\mathcal{V}(RoH)$ the volume occupied by each RoH. Colours indicate the fraction of volume of each RoH as shown in (a).

The spatial location of the $H - \mathcal{K}^{vw}$ nodes is shown in a 3D view in Figure 6.11(a). Highly connected nodes are not scattered in the domain, but they tend to locally group into clusters elongated in the streamwise direction (the longest one with a streamwise extension of about $\Delta x^+ \approx 600$). By following the definition of region provided in this work (namely a spatially-connected set of nodes, as illustrated in Figure 6.8), the $H - \mathcal{K}^{vw}$ nodes shown Figure 6.11(a) are grouped in regions. By doing so, the $H - \mathcal{K}^{vw}$ nodes form in this case 31 regions, which we call regions of hubs (RoHs). It is crucial to recall that nodes in the same RoH are not necessarily all linked to each other; some of them may be linked, but the RoHs merely identify groups of high $\mathcal{K}^{vw}$ nodes belonging to the same spatially-connected region. Such RoHs have different sizes, as illustrated in Figure 6.11(a) where colors indicate the fraction of volume occupied by each RoH, namely $\mathcal{V}_{RoH}/\mathcal{V}_{tot}$. Furthermore, Figure 6.11(b) shows the wall-normal coordinate of the center of mass of each RoH, labelled from 1 to 31 as in Figure 6.11(a). By comparing the y^+ position of the RoHs with the average $\mathcal{K}^{vw}$ values in Figure 6.6(a), it emerges that the presence of the biggest RoHs (around $y^+ \approx 15$, RoHs labelled as 7-12, 24, 26, and $y^+ \approx 120$, RoHs labelled as 15, 16, 22, 23) is the main responsible of the local peak values of $\mathcal{K}^{vw}$.

The effect of the correlation threshold, θ, on the spatial location of the RoHs is shown in Figure 6.12 for $\theta = \{0.8, 0.85, 0.9, 0.95\}$. As θ is changed, $H - \mathcal{K}^{vw}$ nodes close to the wall preserve their overall spatial organization, i.e. they group into x-elongated spatially-connected regions. For $H - \mathcal{K}^{vw}$ nodes close to the center of the channel, a similar scenario is found for $\theta \geq 0.85$, while for $\theta = 0.8$ nodes far from the walls tend to have lower $\mathcal{K}^{vw}$ values (i.e., RoHs disappear) than nodes close to the wall. This is due to the fact that nodes close to the walls are highly involved in teleconnections for low θ values (see also Figure 6.7(a)), thus markedly increasing their centrality in the network with respect to nodes in the channel center.
The elongated shape of the RoHs can be seen as an effect of the mean flow in the streamwise direction. In fact, the presence of streamwise-elongated RoHs is not strongly dependent on the variable selected, since they appear also for the network built on the wall-normal velocity component (as better discussed in the next Section 6.4.2). Moreover, the effect of the mean flow is maintained for sufficiently high correlation thresholds (as shown in Figure 6.12). The mean flow, therefore, induces the characteristic shape of RoHs; indeed, this feature is not detected for high-degree nodes in the correlation network for homogeneous isotropic turbulence [148].

To summarize, the most connected nodes are localized at specific y^+ coordinates and they also tend to cluster in space into elongated structures. The occurrence of similar patterns of RoHs throughout the domain is a remarkable outcome. In fact, one would expect different spatial patterns of $H - \mathcal{K}^{vw}$ nodes at different y^+, because the two-point correlation of the streamwise velocity changes along y^+ (see the average behaviour at different y^+ in Figure B.1(b)). Instead, although the network is based on the two-point correlation, it is able to retain, all at once,

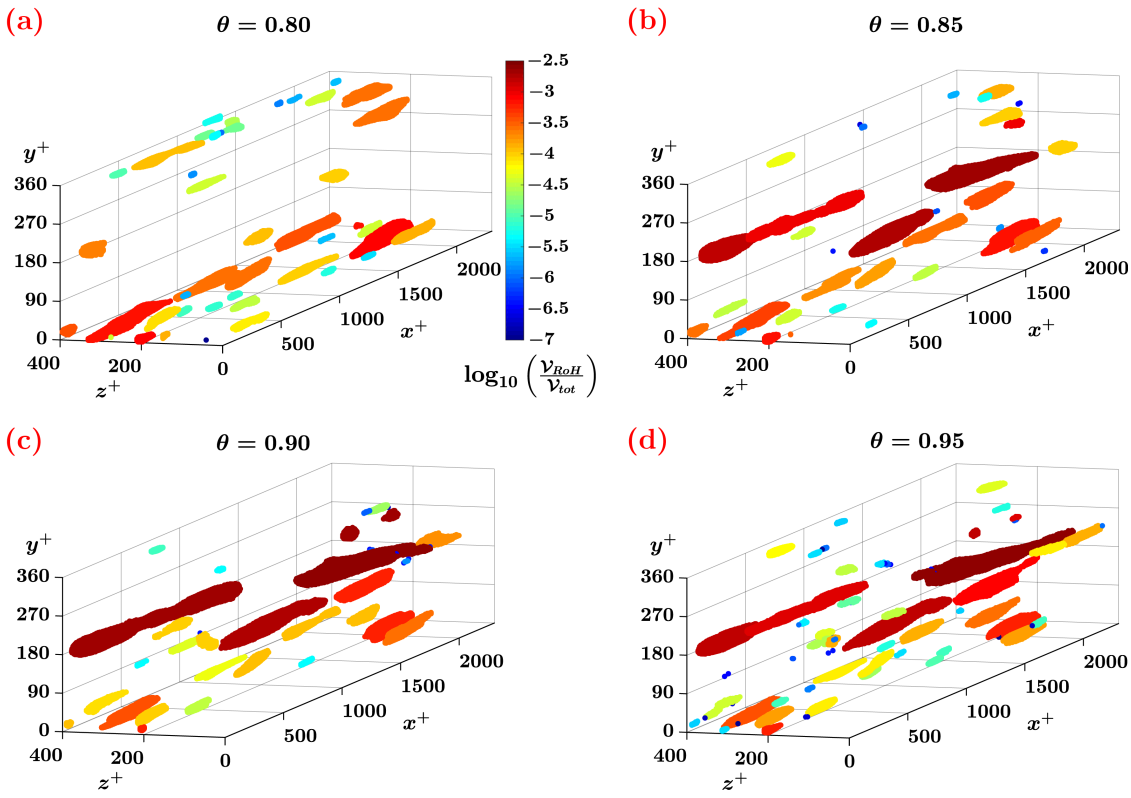


Figure 6.12: 3D views of high-$\mathcal{K}^{vw}$ nodes, namely regions of hubs, RoH for different threshold values, θ. Colour scale refers to the fraction of volume occupied by distinct RoHs, $\mathcal{V}_{RoH}/\mathcal{V}_{tot}$.

the multi-point effects of active links in all directions, for nodes at any y^+. This outcome emphasizes the potential of the complex network approach to enrich the spatial characterization of wall turbulence through a unique framework .

Physical interpretation of teleconnections

The mesoscale analysis carried out so far has evidenced the presence of long-range connections – called teleconnections – between near-wall nodes, and the organization in spatially-connected regions of highly linked nodes (RoHs), both in the near-wall and outer layer. In this Section, we combine these two main results to investigate the first neighbourhood of the most central nodes that also have long-range neighbours. In other terms, we explore the long-range neighbourhoods of nodes in the RoHs, which are only present close to the walls. It is worth to recall that highly-connected nodes (hubs) belonging to the same RoH only satisfy a proximity criterion, but they are not necessarily linked with each other (topological criterion).

To this aim, in Figure 6.13 we show an example of RoH (depicted in black and

corresponding to the ninth RoH in Figure 6.11) and two regions ($\mathcal{R}_{L,1}$ and $\mathcal{R}_{L,2}$, depicted in blue and red, respectively) formed by the union of long-range neighbours of the (black) nodes in the RoH. As can be seen, the regions $\mathcal{R}_{L,1}$ and $\mathcal{R}_{L,2}$ inherit the same elongated shape in the streamwise direction, x, and similar volumes of the corresponding RoH. Such a behaviour is found for all RoHs close to the walls and their long-range neighbours. Therefore, not only first neighbours of single nodes at $y^+ \lesssim 70$ form long-range regions (as already observed in Figure 6.7 and Figure 6.9), but here we find that hubs belonging to the same RoH generate long-range regions (i.e., sets of long-range neighbours that are physically close one to each other). In other words, long-range neighbours of nodes in the same RoH are not scattered in the domain but constitute themselves spatially-connected regions. This outcome extends the meaning of teleconnections from single nodes to spatially-extended regions: in the domain, there are *regions* of highly connected nodes (i.e., the RoHs) that are strongly linked with distant spatially-connected regions. These pairs of *teleconnected regions*, therefore, represent near-wall portions of the domain tightly correlated over time from the streamwise velocity point of view, i.e. spatially-extended regions sharing similar dynamics in time.

This outcome plays a crucial role for the statistical significance of the network, specifically for the link activation. In fact, in networks that are built on the basis of statistical similarities (in this case, the correlation coefficient), spurious links may arise. These kind of connections, despite they represent statistically strong dependences (i.e., high correlations), may be present only because of random coincidences instead of a physical mechanism. To assess the statistical significance of links in spatial networks for climate analysis, *Boers et al.* [140] suggested to retain only links (especially teleconnections) that are part of significant spatial bundles. By exploiting the fact that the system under study is spatially embedded, links that are caused by a physical mechanism should exhibit spatial coherence, in contrast to links caused by random coincidences [140]. Therefore, teleconnected neighbours in the network built from the streamwise velocity can be reliably assumed to be statistical significant, since they tend to cluster in space (namely, to form *bundles* of nodes), thus exhibiting spatial coherence.

Another important element to characterize the neighbourhood of the hubs (or RoHs) is the sign of the correlation of links. For a generic node i, we find that the first neighbours of i belonging to a (short- or long-range) region are either all positively or all negatively correlated with i. In other terms, for any node i in the network, the regions formed by its first neighbours are never partially positively/negatively correlated with i, but always exhibit the same correlation sign. As a consequence, by focusing on long-range links only, each region formed by the union of long-range neighbours of nodes of an RoH has a unique correlation sign with the corresponding RoH. For example, in Figure 6.13, the nodes of the selected RoH (coloured in black) are all positively correlated with the corresponding neighbours in $\mathcal{R}_{L,1}$ (coloured in blue), and all negatively correlated with those in

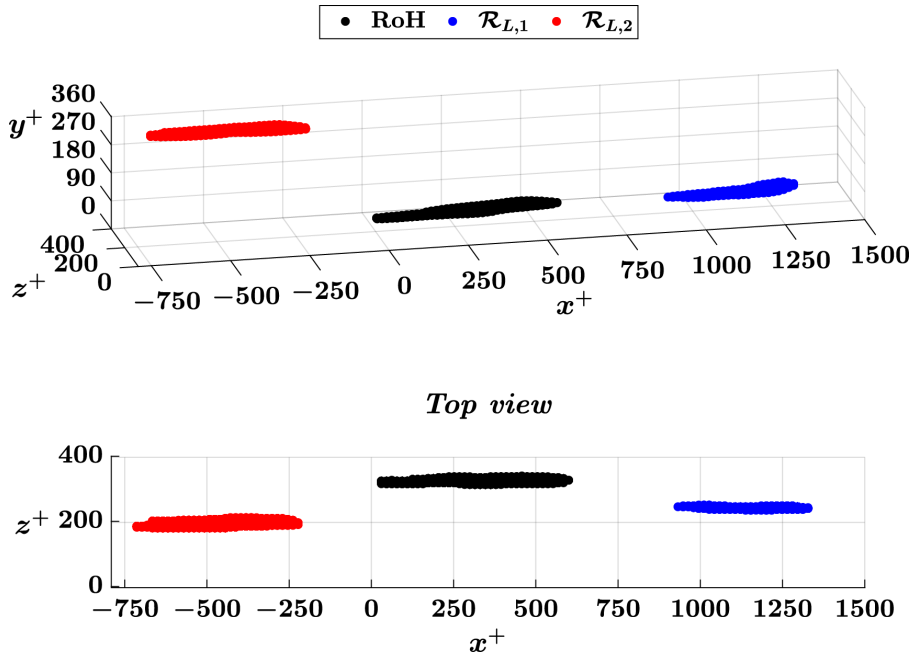


Figure 6.13: 3D views of nodes in an RoH (black) and two $\mathcal{R}_L$ regions formed by long-range neighbours of nodes in the RoH (only a fraction of nodes in the RoH is actually linked to each $\mathcal{R}_L$ region, because these regions are the union, and not the intersection, of the neighbourhoods of the nodes in the RoH). The y^+ values of the center of mass of each $\mathcal{R}_L$ are $y^+ \approx 39$ and $y^+ \approx 353$ (i.e., $y^+ \approx 7$ to the closer wall), while the y^+ of the RoH is $y^+ \approx 11$. The blue and red colors indicate a positive and negative correlation with the RoH, respectively.

$\mathcal{R}_{L,2}$ (coloured in red). Therefore, to summarize, pairs of teleconnected regions – not only pairs of nodes – are found, which correspond to spatially-extended regions of fluid moving with similar streamwise velocity in time, characterized by either positive or negative correlations.

Based on the present findings, we can provide a physical explanation of the teleconnection patterns by relying on the inspection of the streamwise velocity time-series. To this aim, two pairs of neighbours are arbitrarily selected from the regions shown in Figure 6.13: the first pair composed of a node in RoH and a neighbour in $\mathcal{R}_{L,1}$, and the second pair composed of a node in RoH and a neighbour in $\mathcal{R}_{L,2}$. For both pairs, in Figure 6.14 the time-series of the streamwise velocity fluctuations, u', are plotted. It can be noted that the streamwise fluctuations have mostly the same sign for large time-intervals: for the pair involving the node in $\mathcal{R}_{L,1}$ (Figure 6.14(a)), during the time interval $\Delta t \approx |0.77 - 0.12| = 0.65H/u_\tau$, while for the pair involving the node in $\mathcal{R}_{L,2}$ (Figure 6.14(b)), during the time interval $\Delta t_i \approx |0.98 - 0.3| = 0.68H/u_\tau$. This behaviour is typical of high/low speed coherent streaks, that is alternating near-wall regions of positive/negative velocity

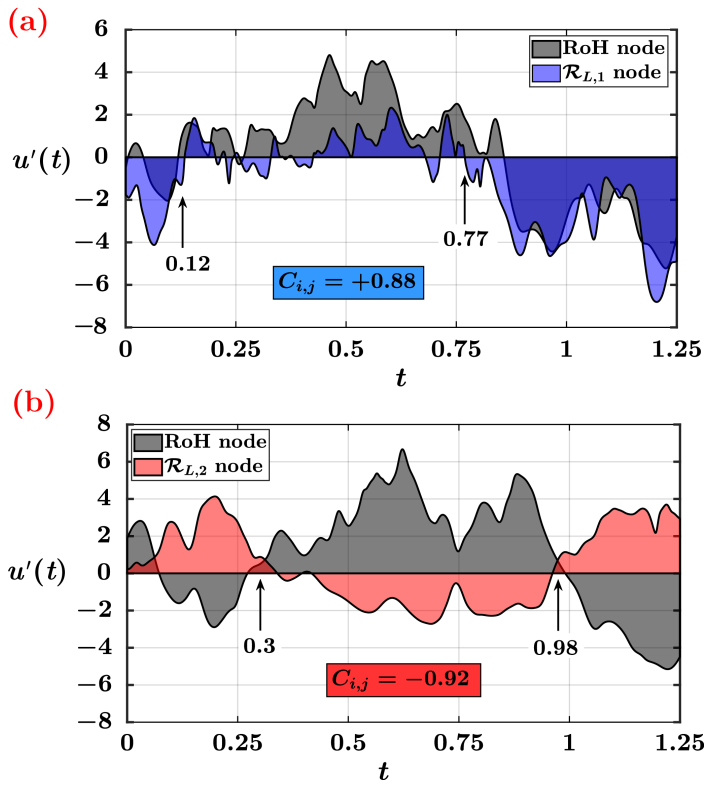


Figure 6.14: Time-series of the streamwise velocity fluctuations, u', of two different pair of nodes extracted from RoHs shown in Figure 6.13 (times are in terms of H/u_τ). The black shaded series correspond to a node in the black RoH of Figure 6.13, while blue and red shaded series correspond to nodes in $\mathcal{R}_{L,1}$ and $\mathcal{R}_{L,2}$, respectively. The values of the correlation coefficients and the temporal limits of the maximum time-interval with the same u' sign for the two pairs of nodes, are also reported.

fluctuations, with an average spanwise separation of $\Delta z^+ \sim 100$ and streamwise lengths $\Delta x^+ \sim 10^3 \div 10^4$ (see Section 2.3.1). By applying the Taylor hypothesis of frozen turbulence [183] and the typical near-wall convective velocity, $U_c^+ \approx 10$ [184], we estimate that the range, Δx^+, corresponding to the time intervals $\Delta t \approx 0.65$, is $\Delta x^+ = (\Delta t \cdot Re_\tau)\, U_c^+ \approx 1200$, which is in agreement with the typical streamwise elongation of the streaks. The presence of near-wall teleconnections with both positive and negative correlation sign can be thus interpreted as an imprint of turbulent coherent structures with time-scale of the order of the temporal window considered. In particular, the complex network approach is able to provide a high level of spatial information (i.e., spatial position, shape and size) of such coherent patterns, thus enriching the spatial characterization of wall-bounded turbulent flows.

6.3.3 Local scale analysis

In this section we exploit the results obtained from the global and mesoscale analyses to further explore the kinematic relations between nodes in the domain. By focusing on representative node pairs, we inquire what are the *correlation paths* in the network between the selected nodes and all the others in the network. In fact, the network built on the correlation coefficient of the velocity time-series is not just a trivial collection of correlated points, but it represents a pattern of (linear) interconnections among nodes. Therefore, the network can be interpreted as a structure of links over which the kinematic information moves throughout the domain. In particular, if a node i is linked to a node j (i.e., by hypothesis, they are strongly correlated) and the node j is linked to a third node k (but i is not linked to k, i.e. $\mathcal{A}_{i,k} = 0$), then the kinematic information indirectly flows from i to k by means of j (i.e., i and k are indirectly linked). The ability to unveil such indirect relations between not-directly-linked nodes is one of the most powerful features of complex networks [83], which is highlighted by the analysis of shortest paths. To this end, we analyse the indirect relations between nodes from the point of view of the shortest paths (see Section 3.3.2) and the cumulative neighbourhoods, Γ_c^n (see Section 3.3.1) of some representative nodes.

High correlation shortest paths

Here we explore the shortest paths between nodes at different wall-normal locations. We recall that a shortest path is the path of minimum cost between two nodes, where the cost represents the *shortest path distance*. Since the direction of the links is not taken into account in this work, the shortest path starting from node i and arriving to node j, is the same as starting from j and arriving to i (i.e., the order of the end-nodes of the shortest path is not relevant). If the links in the network are weighted (i.e., a scalar value is assigned to each link), the shortest path distance corresponds to the minimum value of the sum of the link-weights between two nodes. When the network links are established on the correlation coefficients, an appropriate metric for weighting links and evaluating the shortest path is the function $\mathcal{D}_{i,j} = \sqrt{2(1 - |C_{i,j}|)}$ [185] (the 2 is conventional), which highlights the paths with high (in modulus) correlation values, $C_{i,j}$. More in detail, $\mathcal{D}_{i,j}$ fulfils the three axioms defining a metric, namely $\mathcal{D}_{i,j} = 0$ iff $i = j$ (as $C_{i,i} = 0$), $\mathcal{D}_{i,j} = \mathcal{D}_{j,i}$ (as $C_{i,j}$ is symmetric) and $\mathcal{D}_{i,j} \leq \mathcal{D}_{i,k} + \mathcal{D}_{k,j}$ for any (i, j, k). The expression of $\mathcal{D}_{i,j}$ as a function of the correlation coefficient can be obtained by following the geometrical reasoning in which each time-series of length N can be viewed as a multidimensional vector in a Euclidean N-dimensional space. Hence, by considering two time-series, $u_i^*(t)$ and $u_j^*(t)$, normalized with their mean and standard deviation values and extracted in two spatial locations, i and j, respectively, the square of the distance between them can be expressed as:

$$\mathcal{D}_{i,j}^2 = \sum_{k=1}^{N} \left[u_i^*(t_k) - u_j^*(t_k)\right]^2 = (N-1)\Bigg[\overbrace{\sum_{k=1}^{N}\left(\frac{u_i^*(t_k)^2}{N-1}\right)}^{=1} + \overbrace{\sum_{k=1}^{N}\left(\frac{u_j^*(t_k)^2}{N-1}\right)}^{=1} + \\ - 2\underbrace{\sum_{k=1}^{N}\left(\frac{u_i^*(t_k)\,u_j^*(t_k)}{N-1}\right)}_{=C_{i,j}}\Bigg] = (N-1)\left[2-2C_{i,j}\right] \sim (1-C_{i,j})\,, \tag{6.1}$$

where the first two summations are both equal to 1 (due to series normalization) since they are the variance of $u_i^*(t)$ and $u_j^*(t)$, respectively, while the third summation is the definition of correlation coefficient. In this work, the absolute value of C is taken in order to give the same importance to positive and negative values. Therefore, from the Eq. (6.1) it follows that if two time-series are highly correlated their distance is short (in an N-dimensional space).

We analyse two configurations of shortest path in which links are weighted through $\mathcal{D}_{i,j}$, by selecting either short/long-range links or short-range links only. In the first configuration, indicated as SL, both short- and long-range links are considered (i.e., the network as it was built) and nodes at different y^+ are investigated. Specifically, two pairs of nodes are selected: (i) an end-node of the path close to the wall and the other end-node at the center (see Figure 6.15(a)), and (ii) both the end-nodes at the center of the channel (see Figure 6.15(b)). The shortest paths for the SL configuration are shown in light-blue in Figure 6.15. The second configuration (indicated as S) is also analysed to highlight the effects of the teleconnections in the shortest paths, so that only short-range links are considered (i.e., long-range links are removed). The resulting shortest paths are shown in magenta in Figure 6.15, for the same pairs of end-nodes as the SL configuration.

For the SL and S shortest paths shown in Figure 6.15(a), the total cost (equal to $\sum \mathcal{D}_{i,j}$) is 7.55 and 8.69, while the number of links is 17 and 18, respectively. For the SL and S shortest paths reported in Figure 6.15(b), instead, the total cost is 12.18 and 18.57, while the number of links is equal to 26 and 38, respectively. The difference between S and SL configurations is evident: by including the teleconnections (i.e., the SL configuration) the shortest paths are more complex and involve nodes in the near wall region, as a consequence of the strong connectivity of this part of the domain. Furthermore, the S shortest paths are made up of more links and have a higher total cost (i.e., lower effective correlation) than the SL shortest paths. It is worth noting that the shortest paths from a node at the channel center towards a node close to the wall in the S configuration (see Figure 6.15(a)) tend to be vertically aligned in the wall-normal direction. This outcome suggests that a top-down interaction between the outer layer and the near-wall regions is present, but it is not direct. In other terms, nodes at the channel center and close to the walls are not directly linked (as also emerged from the mesoscale analysis) but they

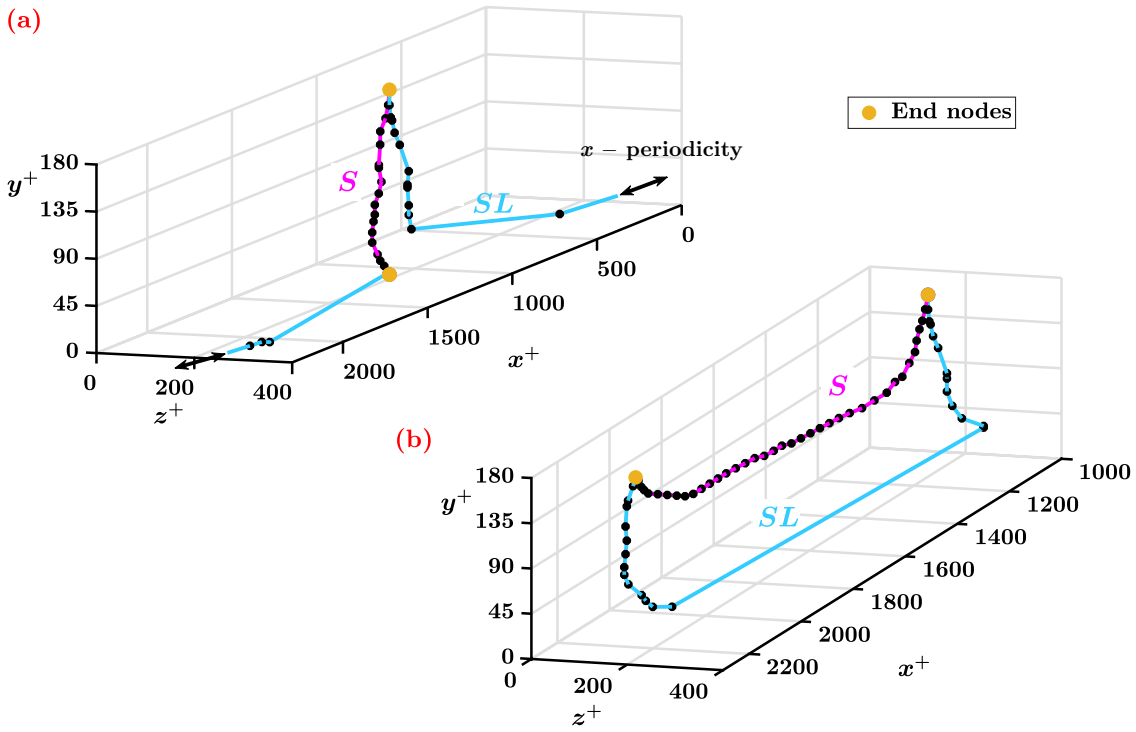


Figure 6.15: Shortest paths in the two configurations: SL, i.e., short/long-range links (coloured in light-blue); S, i.e. only short-range links (coloured in magenta). The end-nodes of the paths are depicted in orange, while intermediate nodes are reported in black. (a) End-nodes are at different y^+, one close to the wall, at coordinates $(x^+, y^+, z^+) = (1139, 3.5, 197)$, and the other at the center of the channel, i.e., $(x^+, y^+, z^+) = (1139, 180, 197)$. The black arrows indicate the periodicity of the domain in the x-direction. (b) Both end-nodes are at the center of the channel, at coordinates $(x^+, y^+, z^+) = (1139, 180, 197)$ and $(x^+, y^+, z^+) = (2262, 180, 197)$.

are indirectly linked via a shortest path that is vertically aligned. This is in line with the idea that the emergence of teleconnections of the streamwise velocity can be interpreted as the footprints of the top-down interactions.

To further underline the importance of teleconnections in the overall topology of the network, we show a particular case of the S configuration, in which the end-nodes of the shortest path are linked with teleconnections. Specifically, the shortest paths are evaluated by selecting as end-nodes a source-node in the near-wall domain (since teleconnections are present for $y^+ \lesssim 70$) and six different teleconnected neighbours of it. These paths are shown, with different colors, in two 3D views in Figure 6.16(a) and Figure 6.16(b): they represent the shortest paths connecting the source-node to its long-range neighbours (and vice-versa) if teleconnections were not present. The removal of the long-range links implies that the kinematic information has to flow through several short-range links, demonstrating that the

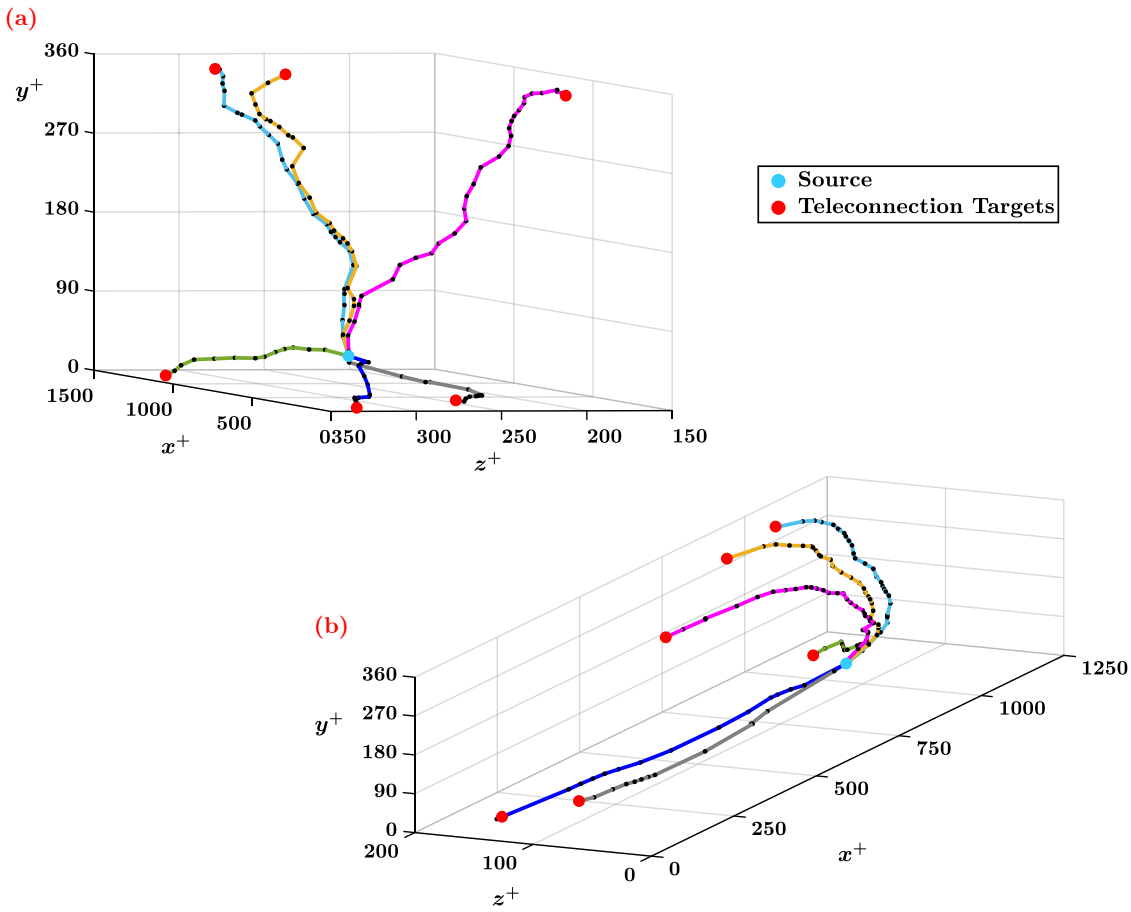


Figure 6.16: Example of shortest paths between a source-node close to the wall (light-blue dot) and six teleconnected neighbours of it, shown through two 3D views. Different colors refer to different targets, while black points indicate intermediate nodes in the paths.

presence of teleconnections provides a bridge for the kinematic information between distant locations in the domain.

Cumulative neighbourhoods analysis

To conclude the analysis at the local-scale level, we explore the behaviour of the cumulative neighbourhoods, $\Gamma_c^n(i)$, for nodes i belonging to three representative pairs of nodes, where each pair comprises a high and a low $\mathcal{K}^{vw}$ node either close or far from the wall. We recall that the n-th neighbourhood of a node, i, is the set of nodes that can be reached from i by crossing at least n different links. Therefore, the n-th cumulative neighbourhood, $\Gamma_c^n(i)$, of i is the union of its first n neighbourhoods (including i). The three representative node pairs are labelled according to their wall-normal coordinate as c, $w1$ and $w2$ indicating that their

coordinates are $y^+ \approx 180$ (channel center) and $y^+ \approx 3.5$ from the closest wall, respectively. Moreover, these nodes are here referred as *source-nodes* as they are the *source* of each cumulative neighbourhood. The behaviour of all other nodes with intermediate y^+ coordinates and $\mathcal{K}^{vw}$ values is not shown as it lies in between the behaviour of nodes in c, $w1$ and $w2$.

It should be pointed out that the behaviour of the cumulative neighbourhoods of a given node, i, provides a synthetic representation of all the shortest paths from i to all other nodes in the network. In fact, since to find the shortest path between two nodes via the optimum algorithm it is required a time-complexity of the order of $O(\mathcal{N}_e)$ [186], for large networks the shortest path evaluation could be computationally demanding. Therefore, the cumulative neighbourhoods provide a simplified but qualitatively informative way to explore the shortest paths between a source-node and the remaining nodes in the network.

Figure 6.17 shows the positions of the cumulative neighbourhoods of the high-$\mathcal{K}^{vw}$ nodes of the pair c (channel center) and $w1$ (i.e., at $y^+ \approx 3.5$), for $n = \{1, 2, 3, 8, 13, 18\}$. It is worth noting that the network is built on a periodic computational domain in the x direction, which markedly affects the three-dimensional views of the Γ_c^n (e.g., see the $n = 8$ panel in Figure 6.17(b)). The resulting spatial expansion of Γ_c^n for nodes at different y^+ appears completely different. By focusing firstly on the node close to the wall (Figure 6.17(a)), the first three neighbourhoods are all close to both walls, revealing the occurrence of teleconnections which are scattered in the (x, z) directions (confirming the outcomes from the mesoscale analysis). In particular, the pattern of these neighbourhoods is very anisotropic, since it is quite elongated in the streamwise direction. Neighbourhoods for $n > 3$ also include nodes close to the walls first (see for example the panel at $n = 8$), starting to incorporate nodes at higher y^+ values and finally approaching the center of the channel. Therefore, in this case, the cumulative neighbourhoods move progressively from the walls to the center of the channel, implying a very high connectivity among nodes close to the walls, where teleconnections play a crucial role. On the other hand, for the high-$\mathcal{K}^{vw}$ node at the channel center (Figure 6.17(b)), the first neighbourhoods are all close to it, they are almost isotropic in the (y, z) directions and a bit elongated along the x direction. Only from the $n = 13$ neighbourhood, the nodes close to the walls start to be included. From this step onward, all the nodes close to the walls are first included in the successive neighbourhood expansion, and only later all the other nodes at higher y^+ are covered. Therefore, the behaviour displayed in Figure 6.17 indicates that, either by considering a node close to the wall or at the center of the channel, the expansion of the cumulative neighbourhoods in the y^+ direction does not occupy the volume of the domain in a monotonic way. In fact, the near-wall regions tend to be included faster, while the central part of the domain is covered afterwards.

It is remarkable to note that the pattern shown in Figure 6.17(b) resembles the SL shortest path shown in light-blue in Figure 6.15(b), in which the path between

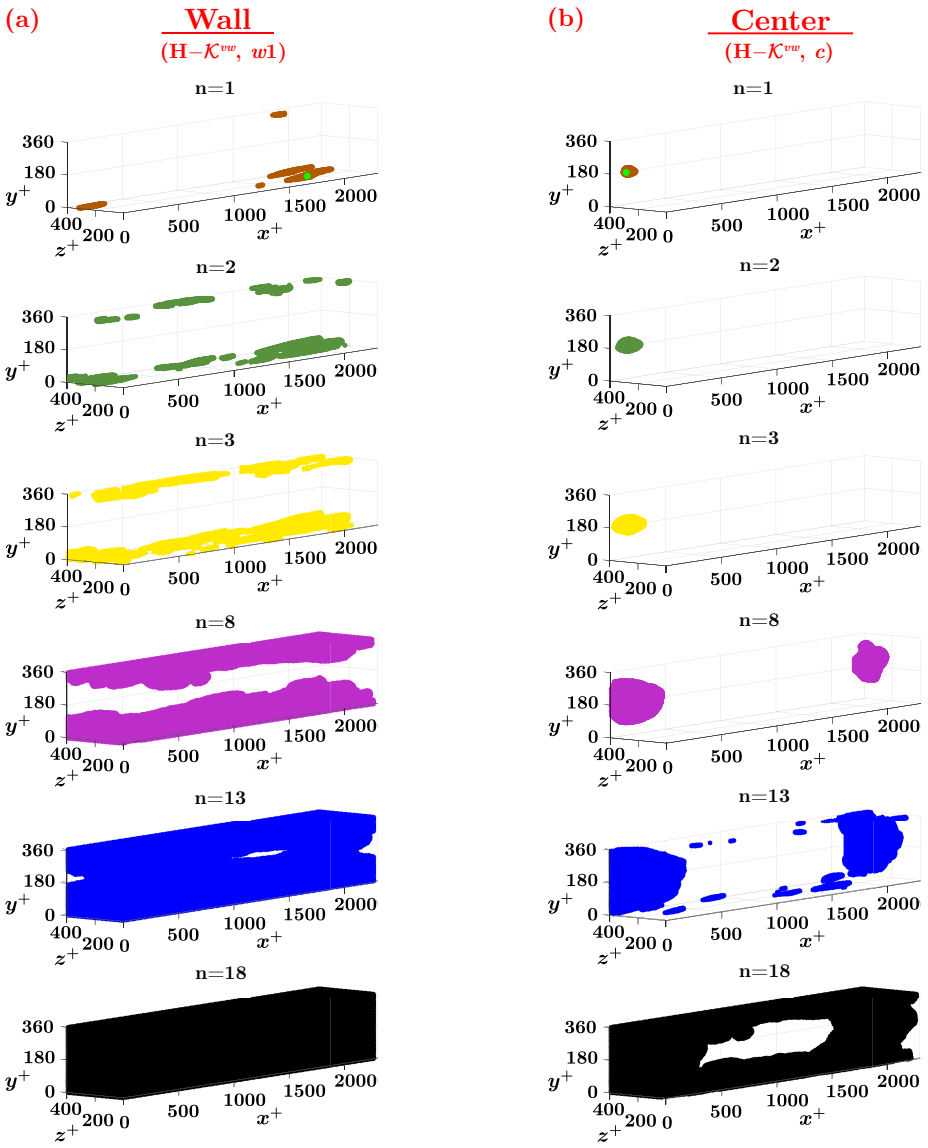


Figure 6.17: (3D views of the n-th cumulative neighborhoods of high-$\mathcal{K}^{vw}$ nodes (marked in green in the $n = 1$ panels). (a) Node close to the wall at $y^+ \simeq 3.5$; (b) node at the center of the channel, $y^+ = 180$.

two nodes at the channel center has to reach the wall region before connecting again to the channel center. This emphasizes the conceptual connection between shortest paths and n-th neighbourhoods, as well as the role of teleconnections as intermediary links for kinematic information flow over long distances, even for pairs of end-nodes at the center of the channel.

The behaviour of the cumulative neighbourhoods that emerge from Figure 6.17

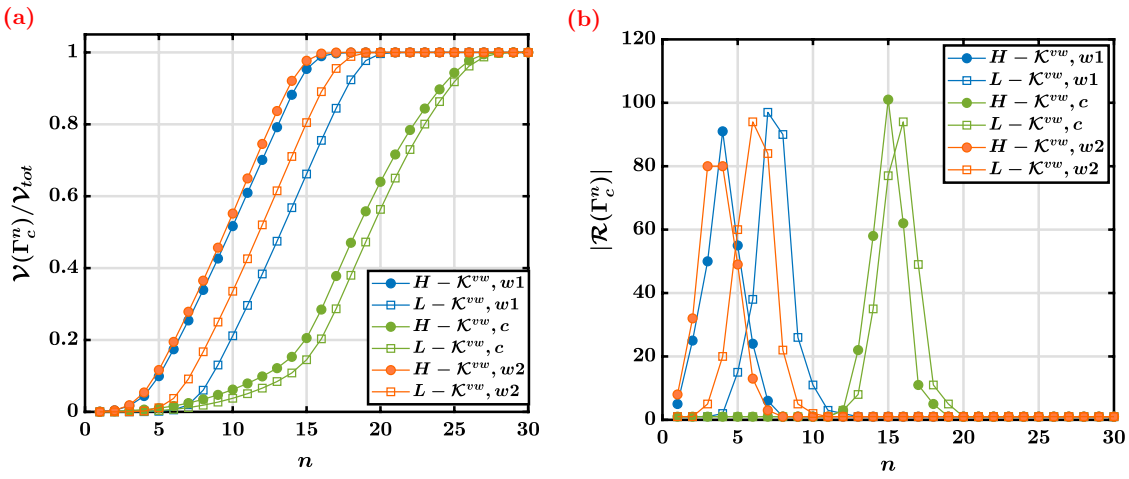


Figure 6.18: (a) Fraction of volume, $\mathcal{V}_N/\mathcal{V}_{tot}$, and (b) number of regions, $|\mathcal{R}(\Gamma_n^c)|$, of the n-th cumulative neighbourhoods for the three selected pairs of nodes ($w1$, $w2$, c).

can be quantified by means of two parameters: the size (expressed as volume fractions) and the number of spatially-connected regions of each cumulative neighbourhood. The fraction of volume, $\mathcal{V}(\Gamma_n^c)/\mathcal{V}_{tot}$, occupied by cumulative neighbourhoods Γ_n^c of the three selected node pairs is shown in Figure 6.18(a), while the number of regions, $|\mathcal{R}(\Gamma_n^c)|$, of each Γ_n^c is displayed in Figure 6.18(b). For both the fraction of volume and the number of regions, pairs of nodes $w1$ and $w2$ show the same overall behaviour as a function of n, which is faster than the one observed for the c pair. Specifically, focusing on the fraction of volume, the expansion of the neighbourhoods for nodes in $w1$ and $w2$ (see blue and orange curves, respectively) is initially much faster than the expansion for the pair c (green curves). However, at intermediate n values (i.e., from around $n = 5$ for the $w1$ and $w2$ pairs, and $n = 15$ for the c pair), the Γ_c^n of all source-nodes tend to growth almost linearly with similar slopes: these ranges of n values correspond to a wall-normal stratified increase of $\mathcal{V}(\Gamma_n^c)/\mathcal{V}_{tot}$ (e.g., see panels for n equal to 8 and 13 in Figure 6.17(a) and $n = 18$ in Figure 6.17(b)). From this range on, the n-th neighbourhoods are composed of almost planar layers of nodes parallel to the wall.
By focusing the number of regions formed by each Γ_n^c, the peaks of $|\mathcal{R}(\Gamma_c^n)|$ occur at low n for the $w1$ and $w2$ pairs, while for the c pair the maximum values of $|\mathcal{R}(\Gamma_c^n)|$ are attained at higher n. Values of $|\mathcal{R}(\Gamma_c^n)|$ greater than one are localized in specific ranges of n associated to the inclusion of nodes close to the walls (with consequent appearance of teleconnections), while for the remaining n values $|\mathcal{R}(\Gamma_c^n)| = 1$ (i.e., only one region). Moreover, within a pair of nodes, the behaviour of the nodes with low-$\mathcal{K}^{vw}$ values is similar but slower than the corresponding node with high-$\mathcal{K}^{vw}$. This is in line with the assortativity plot in Figure 6.5(a) where nodes with low $\mathcal{K}^{vw}$ values are more likely to be linked to nodes with similar $\mathcal{K}^{vw}$ and some more

steps are required to reach nodes with high-$\mathcal{K}^{vw}$ values, which are connected to a larger fraction of the domain.

The analysis of the Γ_c^n neighbourhoods provides insights into the kinematic information flow, evidencing that: (i) nodes in the near-wall regions (indicatively, $y^+ \lesssim 70$) are strongly inter-connected for low n values, creating anisotropic textures of teleconnections, that result in a very effective kinematic information spreading; (ii) nodes around the center of the channel (indicatively, $y^+ \gtrsim 70$) display localized high-correlation patterns for low n values, similar to those extracted in homogeneous isotropic turbulence [148]; (iii) high $\mathcal{K}^{vw}$ nodes are the most central in the network, not only relative to the first neighbourhood but also in relation with the whole network.

To summarize, the local scale analysis provides a detailed description of the topological and kinematic relations between different physical locations in the domain. In particular, the investigation of Γ_c^n and the shortest paths reveals the full potential of the network in capturing the spatial information related to the patterns of indirect interactions and teleconnections, which is possible only thanks to a network approach.

6.4 Complementary results

In this Section we report results that are complementary to the main analysis described in the previous Section 6.3. We explore how different parameters and variables affect the correlation-based network, by highlighting similarities or differences with respect to the principal analysis. In particular, the main features of the network built on the streamwise velocity for a higher Reynolds number and different time windows is explored in Section 6.4.1. The network built on the wall-normal velocity component, instead, is investigated in Section 6.4.2.

6.4.1 Parametric analysis of the u-based network

Additional results are here provided for the spatial network built on the streamwise velocity component. First, the effects of a higher Reynolds number is investigated by exploiting a DNS at $Re_\tau = 590$ for $T^+ \approx 370$ (details of the simulation are included in Appendix B.4). The spatial network is built by following the same procedure as the case for $Re_\tau = 180$. In particular, the correlation coefficients are filtered by using $\theta = 0.85$ as in the main analysis reported in Section 6.3.

Figure 6.19(a) shows the cumulative probability distribution of the volume-weighted connectivity (black curve). As in the network for $Re_\tau = 180$ (Figure 6.4(a)), the cumulative distribution follows an exponential decay with a cut-off for large $\mathcal{K}^{vw}$ values. This suggest the presence of hubs in the network, similarly

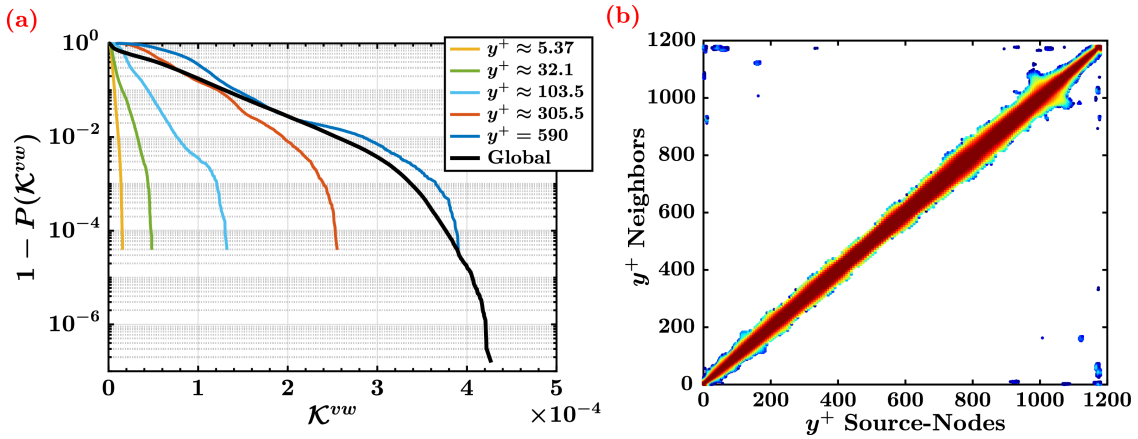


Figure 6.19: (a) Cumulative $\mathcal{K}^{vw}$ distribution, $1 - P(\mathcal{K}^{vw})$, for the network built on the streamwise velocity at $Re_\tau = 590$. The behaviour of $P(\mathcal{K}^{vw})$ is shown for the global network (black curve) and for five representative y^+ coordinates. (b) Probability (in percentage and in $\log_{10}$ scale) that a source-node at a given y^+ is linked to a neighbour at another y^+ value for the network at $Re_\tau = 590$.

to the case at low Reynolds number. In order to explore the presence of teleconnections, the probability that an arbitrary source-node at a fixed y^+ plane has a neighbour at others y^+ values is shown in Figure 6.19(b). For the selected time window, T^+, the network built at $Re_\tau = 590$ display fewer teleconnections between near-wall regions than the $Re_\tau = 180$ case (see Figure 6.7(b)). However, for the higher Reynolds number, non-zero probabilities are found for larger y^+ values (up to $y^+ \approx 400$), while for $Re_\tau = 180$ teleconnections are limited to $y^+ \lesssim 70$.

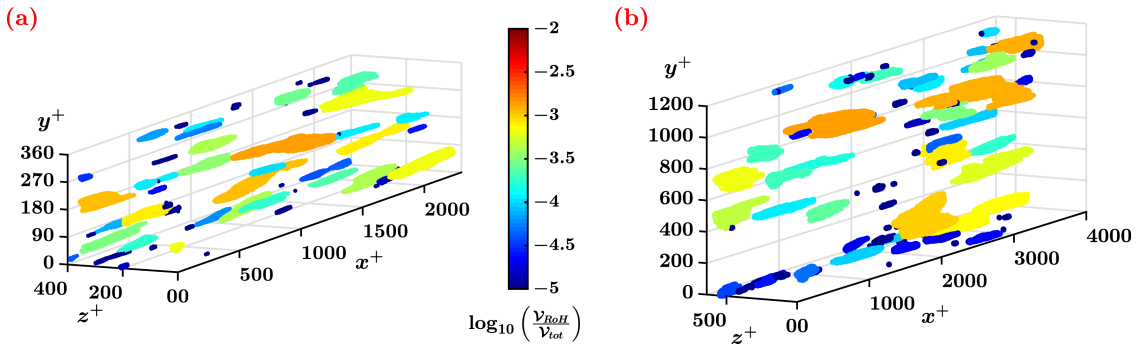


Figure 6.20: Comparison of the regions of hubs (RoHs) for the networks built at (a) $Re_\tau = 180$ and (b) $Re_\tau = 590$. RoHs are defined based on the 99th percentile of the $\mathcal{K}^{vw}$ probability distribution conditioned to the y^+ coordinates of nodes. Colours indicate the fraction of volume of each RoH.

The presence of regions of hubs (RoHs) for the network built at $Re_\tau = 590$ is also investigated. However, a direct comparison between RoHs at different Reynolds

numbers cannot be done, because the domain size in wall-units increases in the wall-normal direction proportionally to the Reynolds number. In fact, since the outer layer extension increases for higher Re_τ, nodes that are far from the wall are more likely to link with other nodes in the outer layer. As a consequence, nodes located at high y^+ values represent cells of the domain with large volume, thus significantly affecting their $\mathcal{K}^{vw}$ value. In other terms, a node at the channel center for $Re_\tau = 590$ is much more likely to have a high $\mathcal{K}^{vw}$ value than a node at the channel center for $Re_\tau = 180$. To overcome this issue, hubs in the network are then re-defined as the most connected nodes (i.e., satisfying the condition $1 - P(\mathcal{K}^{vw}) \leq 10^{-2}$) *at each* y^+ *coordinate*. Namely, the importance of a node, i, in the network is conditioned to the node wall-normal coordinate, y_i^+. Figure 6.19(a) shows the cumulative probability distributions of $\mathcal{K}^{vw}$ at five representative y^+ coordinates for $Re_\tau = 590$. By taking the 99th percentile of each probability distribution at each y^+, we obtain the hubs in the network as shown in Figure 6.20. It can be observed that, by conditioning the hub definition to the y^+ coordinates, spatially-connected regions of hubs are present both for the network built at $Re_\tau = 180$ (Figure 6.20(a)) and $Re_\tau = 590$ (Figure 6.20(b)), and they are still elongated in the streamwise direction as an effect of the mean flow advection.

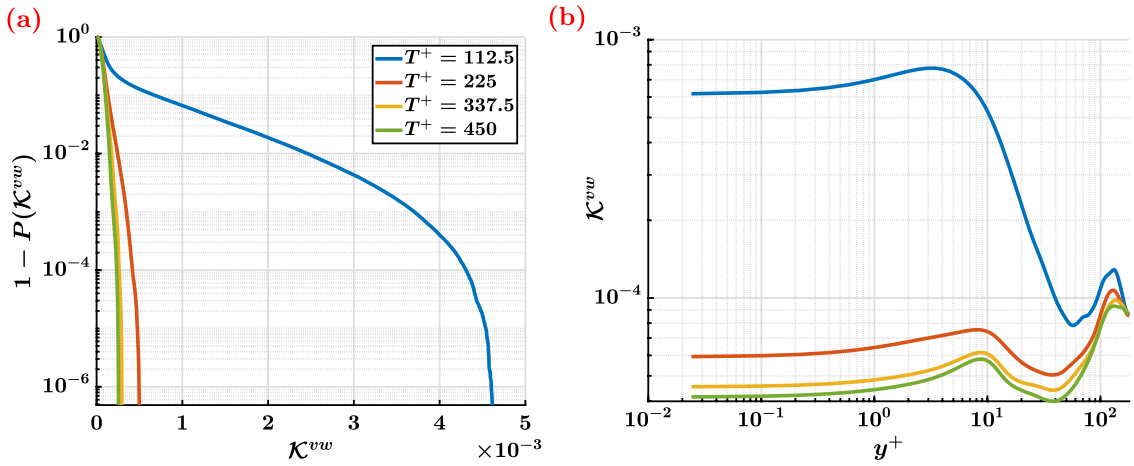


Figure 6.21: (a) Cumulative $\mathcal{K}^{vw}$ distribution for the network built on the streamwise velocity at $Re_\tau = 180$, for different temporal windows, T^+. (b) Mean values of the volume-weighted connectivity as a function of y^+ for different T^+.

To conclude the analysis of the spatial networks built on streamwise velocity component, the effect of the temporal window, T^+, is studied. With this aim, networks are built for three T^+ values different than the reference case at $T^+ = 225$, namely $T^+ = \{112.5, 337.5, 450\}$ corresponding to half, three halves and double the reference case. Figure 6.21(a) shows the cumulative probability distribution for the four T^+ here focused, while the average $\mathcal{K}^{vw}$ as a function of y^+ is reported in

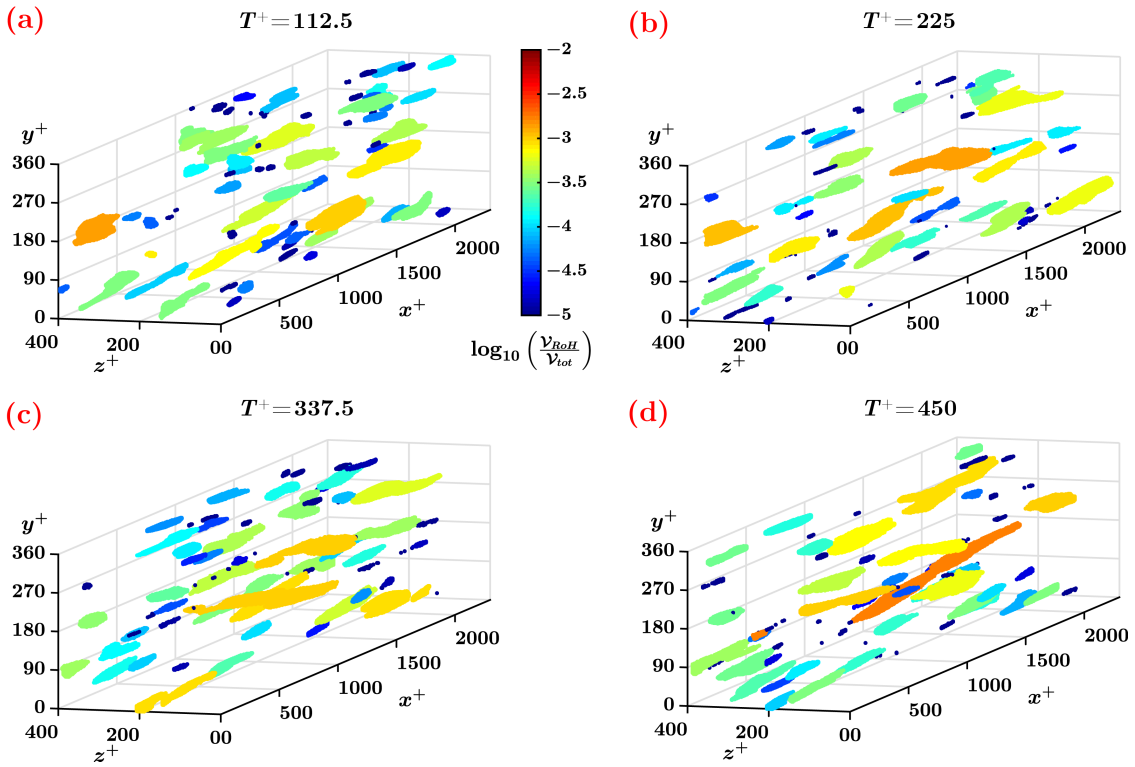


Figure 6.22: 3D views of the RoHs for different temporal windows for the network built at $Re_\tau = 180$. (a) $T^+ = 112.5$; (b) $T^+ = 225$; (c) $T^+ = 337.5$; (d) $T^+ = 450$. RoHs are defined based on the 99th percentile of the $\mathcal{K}^{vw}$ probability distribution conditioned to the y^+ coordinates of nodes. Colours indicate the fraction of volume of each RoH.

Figure 6.21(b). As T^+ decreases, the probability to find nodes with high volume-weighted connectivity increases (Figure 6.21(a)), because it is more likely to find portions of the domain that are highly correlated for shorter temporal windows. The parts of the domain that benefit of the increase of $\mathcal{K}^{vw}$ are those close to the walls: as illustrated in Figure 6.21(b), in fact, the average $\mathcal{K}^{vw}$ strongly increases in the proximity of the wall as T^+ decreases. This in turn implies that teleconnections are more likely to appear if the temporal window considered decreases.

Finally, the spatial location of the hubs of the networks built at different T^+ is shown in Figure 6.22. As for the networks at different Reynolds numbers, the high-$\mathcal{K}^{vw}$ nodes are selected by taking the 99th percentile of the cumulative distributions evaluated at each y^+ coordinate. This choice is motivated by the large differences in the values of $\mathcal{K}^{vw}$ among different T^+ cases, as highlighted by the tails of the distributions in Figure 6.21(a) and the average values in Figure 6.21(b). As previously, spatially-connected regions of hubs elongated in the streamwise directions are found at different wall-normal locations. Additionally, it appears from

Figure 6.22 that RoHs tend to occupy similar portions of the domain as T^+ changes, suggesting that RoHs are the result of a dynamics of flow that persists in time in those locations.

6.4.2 Results for the wall-normal velocity, v

This Section reports some results concerning the network based on the wall-normal velocity, v, with $\theta = \{0.8, 0.85, 0.90, 0.95\}$ as in the main analysis for the streamwise velocity. The wall-normal velocity, besides the streamwise component, is one of the most significant variables to characterize a turbulent channel flow as it is the velocity component corresponding to the inhomogeneous direction.

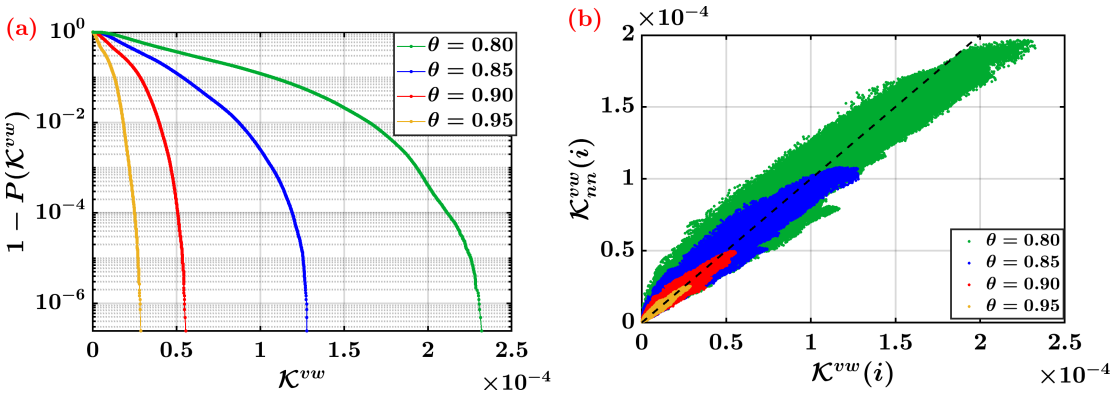


Figure 6.23: Network built on the wall-normal velocity, v. (a) Cumulative $\mathcal{K}^{vw}$ distribution and (b) weighted average nearest neighbours assortativity measure, $\mathcal{K}^{vw}_{nn}(i)$, as a function of $\mathcal{K}^{vw}$, for different thresholds θ.

From a global point of view, sharp decays of $1 - P(\mathcal{K}^{vw})$ are found for different θ values in the network based on v as shown in Figure 6.23(a). Therefore, high values of $\mathcal{K}^{vw}$ are extremely rare in the network based on the wall-normal velocity component. Moreover, as shown in Figure 6.23(b), the network built on v is strongly assortative, i.e., the nodes and their neighbours are close in space and share similar neighbourhoods in terms of spatial extension (i.e., similar values of $\mathcal{K}^{vw}$). Therefore, the global features of the network built on v are similar to those of the network built on u, but more pronounced.

From a mesoscale point of view, by focusing on the threshold $\theta = 0.85$, the average and standard deviation values of $\mathcal{K}^{vw}$ as a function of y^+ are reported in Figure 6.24(a). Differently from the network built on the streamwise velocity component, u, most of the hubs in the network for v are located far from the walls, suggesting an almost absence of near-wall teleconnections. This is indeed confirmed by the probability that an arbitrary source-node has a neighbour at another y^+ value. As shown in Figure 6.24(b), most of the nodes connect with nodes close to them (diagonal part of the plot), and only few points very close to the wall

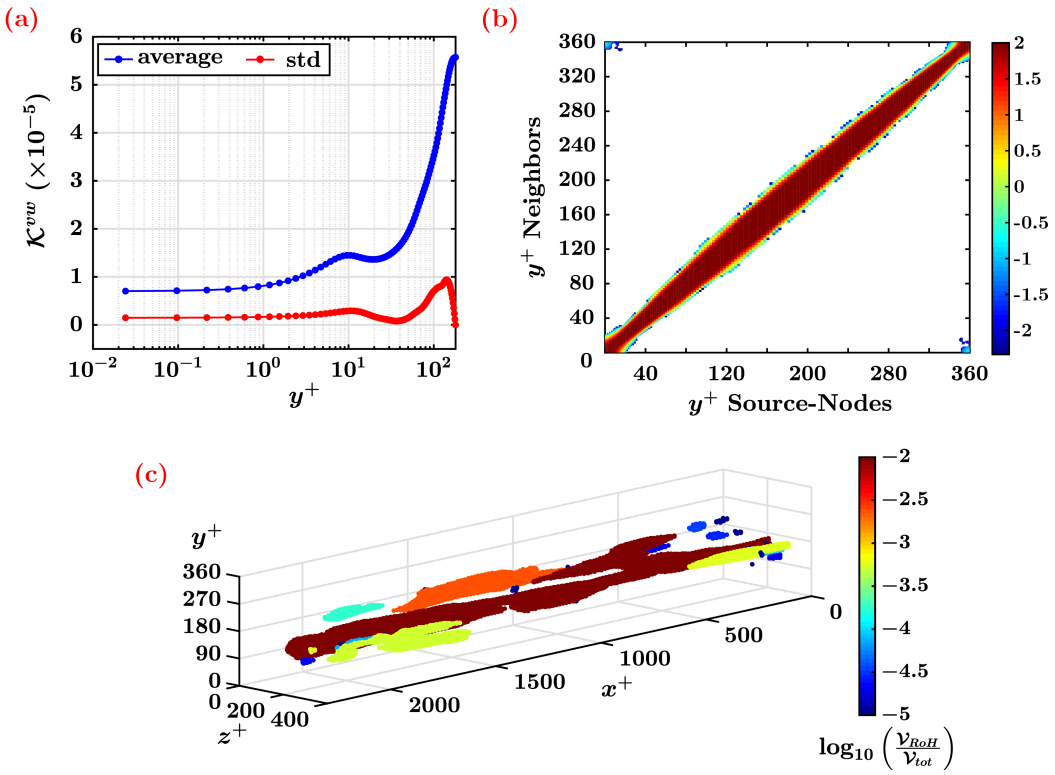


Figure 6.24: Network built on the wall-normal velocity, v. (a) Mean and standard deviation values (averaged over the two homogeneous directions) of the volume-weighted connectivity, $\mathcal{K}^{vw}$, as a function of y^+ for $\theta = 0.85$. (b) Probability (in percentage and in $\log_{10}$ scale) that a source-node at a given y^+ is linked to a neighbour at another y^+ value. (c) 3D view of RoHs (defined from the 99th percentile of the global $P(\mathcal{K}^{vw})$ shown in panel (a)). Colour scale refers to the fraction of volume occupied by distinct RoHs, $\mathcal{V}_{RoH}/\mathcal{V}_{tot}$.

have teleconnected neighbours close to the other wall. The substantial absence of teleconnections, therefore, makes the network of v even more assortative than the network of u. Finally, the spatial location of the $H-\mathcal{K}^{vw}$ nodes of the network of v is shown in a 3D view in Figure 6.24(c). As for the network of u, $H-\mathcal{K}^{vw}$ nodes are not scattered in the domain, but they tend to locally group into clusters elongated in the streamwise direction, as an effect of the streamwise advection. However, differently from the network of u, $H-\mathcal{K}^{vw}$ nodes only occur around the center of the channel for the network based on v. This happens because teleconnections enhance the centrality of nodes close to the walls, and this behaviour is magnified in the network based on u rather than in the network on v.

6.5 Summary and future outlooks

Complex networks are exploited to study of a fully-developed turbulent channel flow at three levels, namely global scale (i.e., considering all nodes, without any distinction), mesoscale (i.e., dealing with groups of nodes), and local scale (i.e., focusing on single nodes). We extend the concept of *area-weighted connectivity* typical of climate networks to a *volume-weighted connectivity*, by constructing a three-dimensional spatial network. In this way, nodes represent fractions of volume of the physical domain, while the correlation coefficient is filtered to activate links, retaining the highest values (in modulus). By doing so, the strongest kinematic linear inter-relations are highlighted.

By investigating the network at different levels, we find the presence of hubs (i.e., nodes highly connected to other parts of the domain), which tend to be localized (on average) at specific wall-normal coordinates, both close to the walls and around the center of the channel. In more detail, hubs tend to cluster into streamwise-elongated regions, RoHs, for both the streamwise and wall-normal velocities and for different Reynolds numbers. However, only in the network built on the streamwise velocity the RoHs appear both close to the walls and in the channel center, while in the network built on the wall-normal velocity they are confined around the channel center. This difference is due to the presence of inter-wall and intra-wall long-range links in all directions in the network for the streamwise velocity. We refer to these long-range links as teleconnections, which reveal an analogous response of the near-wall velocity field to large-scale motions. Additionally, the teleconnected neighbours of nodes in the same RoHs tend to group into spatially-connected regions (similar to the corresponding RoHs). Therefore, RoHs and the corresponding teleconnected regions constitute strongly correlated near-wall parts of the domain, that turn out to be related to the persistence of streamwise velocity streaks, namely near-wall coherent structures.
Finally, to highlight the different ways in which the kinematic information flow in the domain, we investigate the behaviour of the successive neighbourhoods and shortest paths. The presence or absence or teleconnections plays a crucial role in the information flow through shortest paths, since teleconnections provide a shortcut to the kinematic information between distant locations, even between pair of nodes at the channel center.

The proposed network-based approach provides a versatile and powerful framework to study complex systems as turbulent flows in a Eulerian viewpoint, especially in the presence of inhomogeneities. Through the application of the network formalism, a different perspective on wall-bounded turbulent flows is introduced, in which the spatial information is preserved and enriched by the multi-point effects of active links in all directions. In fact, teleconnections between distant near-wall regions have been localized and associated with the temporal persistence of coherent patterns in a straightforward way; this operation may turn out to be a complicated

task when other techniques are adopted.

Future research will be focused on deepening the Reynolds number effects on the network structure, in particular for the emergence of teleconnections. More insights into the interpretation of network hubs (and their spatial organization) as well as velocity teleconnections surely deserve to be explored. To this purpose, the interplay between other turbulent quantities could be studied, for instance, by focusing on the energy or vorticity fields. In this way, similarly to the successful employment in climate analysis, spatial networks can be employed to support classical statistics in the understanding of the top-down mechanisms between large scale motions and near-wall flow dynamics, especially through the characterization of teleconnections. Furthermore, spatial networks can be used for the identification of coherent structures or regions with high turbulence intensity, so that complex networks could also serve as a diagnostic tool for control strategies in practical applications involving wall-bounded turbulence.

Chapter 7

Lagrangian mixing in wall turbulence: a network perspective

Some of the contents presented in this Chapter have been previously published. Below are provided the references:

G. Iacobello, S. Scarsoglio, J.G.M. Kuerten, and L. Ridolfi. "Lagrangian network analysis of turbulent mixing". In: *Journal of Fluid Mechanics* 865 (2019), pp. 546–562. DOI: 10.1017/jfm.2019.79.

7.1 Motivation

The Lagrangian viewpoint of turbulent mixing in which particles are tracked in the flow, has turned out to be better suited than the Eulerian view in many cases, due to the natural way to approach particle dispersion by analysing particle trajectories. By exploiting the Lagrangian viewpoint, the time evolution of pair- and multi-particles has usually been explored in terms of geometrical features, such as pairwise mean-square separation or multi-particle shape evolution (see Section 2.3.2). In the context of Lagrangian passive scalar dispersion, most of the studies have focused on homogeneous isotropic turbulence [7, 55, 56], with few applications on relative dispersion on anisotropic and inhomogeneous turbulence, such as wall-bounded turbulence [62, 64–66].

In this Chapter, motivated by the need to provide a better characterization of Lagrangian mixing in wall-bounded turbulence, a complex networks analysis of mixing is carried out. In more detail, a turbulent channel flow is investigated in a Lagrangian way, with the aim to extend the level of information of classical statistics (such as the pairwise mean-square separation), thus providing a geometrical representation of the particle dynamics. In so doing, it is possible to shed light on the interplay between the mean flow advection and wall-normal turbulent mixing

on particle dynamics, thus highlighting the key role of the spatial inhomogeneity.

Although network-based analyses of fluid flows in a Lagrangian frame have recently been carried out, they mainly focuses on fluid transport and coherent structure identification. Moreover, the linking criterion has usually been based on a similarity measure between particle trajectories evaluated in a given time interval, so that the temporal details of particle trajectories do not explicitly emerge (see Section 4.4). Differently, in this work the spatio-temporal evolution of particles is captured at each time step, providing a rich and detailed time-dependent picture of turbulent mixing. Therefore, the time-varying network formulation here proposed intrinsically highlights the temporal development of particle dynamics due to the turbulent motion.

The Chapter is organized as follows. Section 7.2 reports the methodology details, by describing the turbulent channel flow setup, the initial particle arrangement, as well as the details of complex network construction. Results are shown in Section 7.3. The main features of the network structure and turbulent mixing characterization are shown in Section 7.3.1 for an unweighted particle interaction. Network metrics are discussed in Section 7.3.2, by highlighting the network structure obtained by introducing a pairwise particle interaction strength. Finally, a resume of the results and future perspectives are delineated in Section 7.4.

7.2 Time-varying network representation of particle dynamics

In order to build a complex network based on particle trajectories, a DNS of a fully-developed (incompressible) turbulent channel flow at $Re_\tau = 950$ is exploited (for more details on the simulation, see Appendix B.5). A set of 100×100 fluid particles is arranged as a uniformly distributed grid in the plane (y^+, z^+) at $x^+ = 0$ at the initial time (i.e., the time in which particle are seeded in the domain). Fluid particles are seeded in the domain after the statistically stationary condition of the flow is reached. The time-dependent particle positions are then extracted by integrating the local (Eulerian) fluid velocity field (see Appendix B.5). Particles are tracked for a time interval equal to $T^+ = 15200$, with a time-step equal to $\Delta t^+ = 4.75$. The total simulation time in wall units, T^+, is set greater than $14200\nu/u_\tau^2$, which is the wall-normal (Eulerian) mixing time, namely the integral time-scale after which the Taylor dispersion analysis is asymptotically valid [188]. The wall-normal (Eulerian) mixing time equals, by definition, H^2/ϵ_y [189], where $\epsilon_y = 0.067\, H\, u_\tau$ is the Elder's vertical mixing coefficient [190]. The initial particle arrangement is displayed in Figure 7.1(a), while Figure 7.1(b)-(d) show the particle positions at three representative times. In Figure 7.1, coordinate-axes are not in scale and particles are coloured according to their initial y^+ values. It is worth

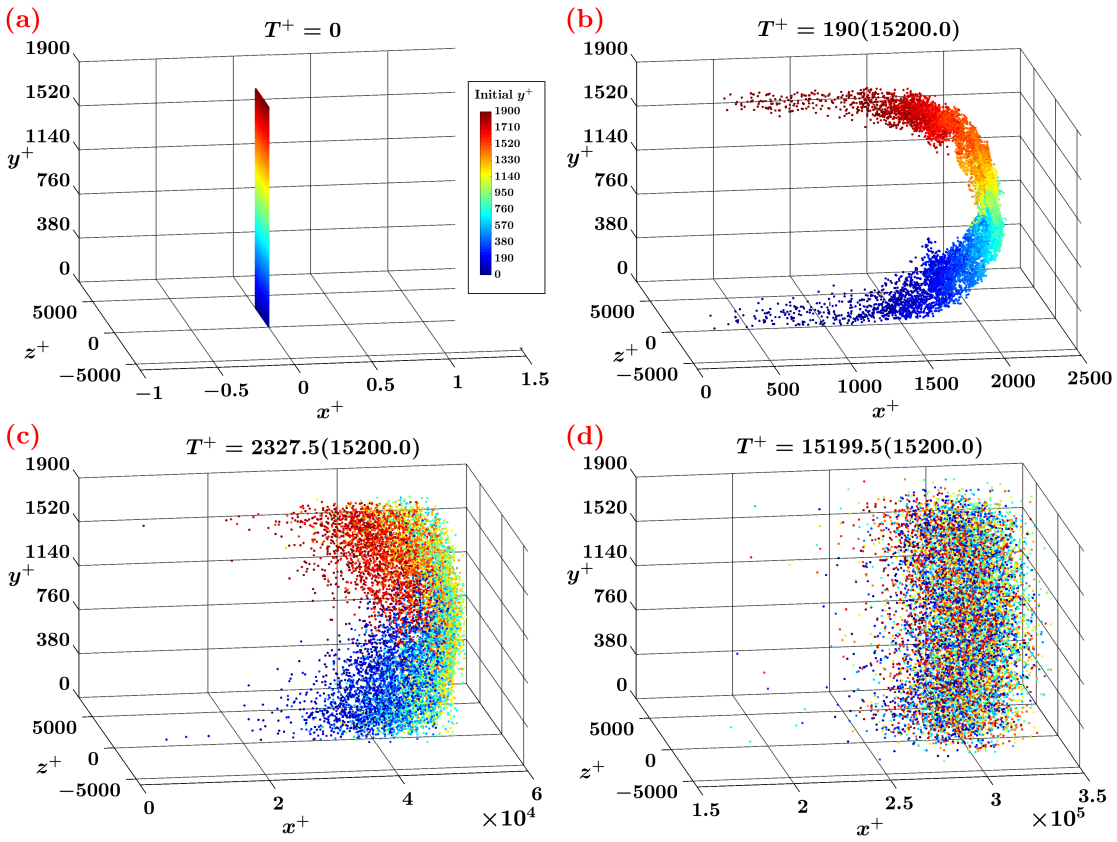


Figure 7.1: 3D views of particle positions at four characteristic times, $t^+ = \{0, 190, 2327.5, 15199.5\}$. Particles are coloured according to their initial wall-normal coordinate, y^+, as illustrated in the inset of panel (a).

noting the different particle arrangements at short times (Figure 7.1(b)) and long times (Figure 7.1(d)).

Particle connection description. To investigate the turbulent dispersion via the network formalism, the interactions between fluid particles are determined based on mutual spatial proximity (as discussed in Section 4.4.2, see Figure 4.5). Namely, two particles, g and q, are connected at a given time if g lies inside the reference ellipsoid centred at q and *vice versa* (by symmetry). Therefore, a connection is established between two particles if they come sufficiently close in space during their motion. This concept is illustrated in Figure 7.2, which shows a sketch of the initial particle configuration as well as the temporal evolution of a particle along its trajectory (depicted as a green dot and dashed line, respectively), and the connections with other particles enclosed in its reference ellipsoid.
We also recall that the ellipsoid semi-axes, a_χ, are set proportional to the average pairwise distances between particles, $\overline{d_\chi(t)}$, in each Cartesian direction $\chi = \{x, y, z\}$

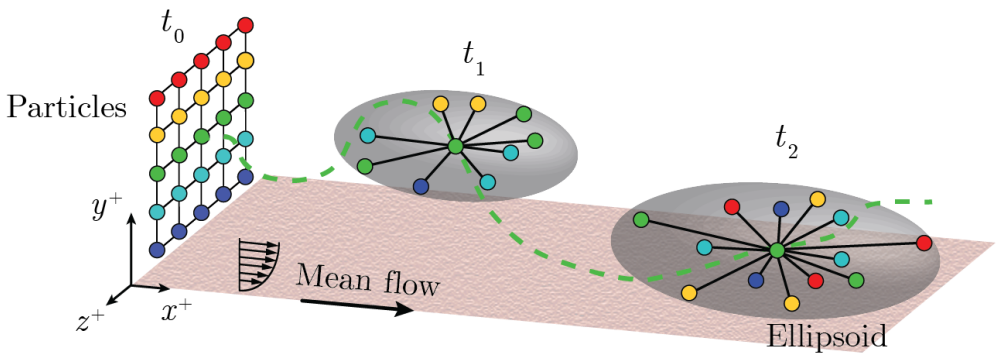


Figure 7.2: Sketch of the setup for the Lagrangian network based analysis. Coloured spheres represent particles, which are initially released from a uniformly spaced grid at $x^+ = 0$. A reference particle is highlighted and reference ellipsoid are depicted as a dashed green line and a grey-shaded surface, respectively. Connections between particles are illustrated as black lines.

(see Figure B.3). Each ellipsoid is determined by means of its semi-axis lengths, $\{a_x, a_y, a_z\}$, representing spatial scales of turbulent motion along each Cartesian direction. Accordingly, if the Euclidean distance between a pair of particles is less than a_χ in each direction $\chi = \{x, y, z\}$ (namely they are connected), then the two particles share turbulent length scales greater (or equal) than a_χ. The choice of a_χ is generally a non-trivial task which depends on the specific problem under study, such as the presence of inhomogeneities in the flow or if particles are involved in chemical or biological process (in which specific interactions occur when particles are sufficiently close). When characteristic scales are not known *a priori*, the issue of setting a_χ – namely to assess the typical length scales in the flow – can be faced, for instance, by relying on turbulence spectra, correlation functions, as well as coherent structure identification techniques. Here, as a reference case, the factor of proportionality, α, between a_χ and $\overline{d_\chi(t)}$ is set as a constant value equal to $\alpha = 0.5$ (so that the ellipsoid size in each Cartesian direction is equal to the average distance). The sensitivity of the network results on α (namely on the ellipsoid size) is also investigated in the range $\alpha \in (0,1)$, showing similar results for different α values (see Section 7.3.1). It should be noted that, although α may also explicitly depend on t^+, we consider α constant in time in order to focus only on the temporal dependency of the semi-axis lengths a_χ.

Particle grouping and network building. Since the focus is on the wall-normal mixing process, particles are grouped into $N_l = 100$ wall-normal levels, l_i. Each level includes a spanwise-row of fluid particles at the initial time (e.g., a level corresponds to a row of particles with the same color in Figure 7.1(a) or Figure 7.2). As a consequence, each level l_i is univocally associated to the y^+ value of particles

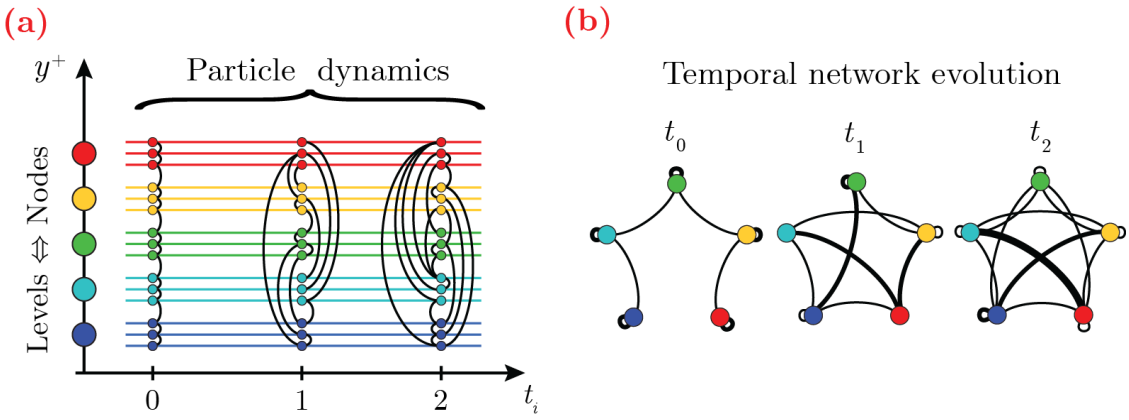


Figure 7.3: (a) Example of particle dynamics depicted as a contact sequence, and (b) temporal evolution of the corresponding networks (the link thickness is proportional to its weight). For both panels, colors from red to blue highlight different starting levels (i.e., different initial y^+ values).

at $t^+ = 0$. The 10^4 particles seeded in the domain are hence arranged in $N_l = 100$ spanwise rows, each one comprising 100 particles separated by $\Delta z^+ = 29.85$. The particles closest to the walls (i.e., l_1 and l_{100}) are located at $y^+ = 9.5$, while the remaining ones are separated by $\Delta y^+ = 19$. The initial particle grouping into levels naturally emerges here as the presence of the walls introduces an inhomogeneous direction, y^+. It should be noted that, differently from previous studies dealing with multi-particle dispersion in which groups of (usually, three or four) particles are "connected" *a priori* and their evolving shape is investigated over time, here connections between particles are dynamically activated over time within a fixed reference ellipsoid.

The spatial proximity approach applied to particle dynamics results in a non-trivial time-sequence of connections, because of the different trajectories followed by particles due to turbulent motion. An example of a geometrical representation of the particle dynamics is shown in Figure 7.3(a) as a contact sequence: each connection between pairs of particles (depicted as small coloured circles) is indicated as a black arc, meaning that particles are sufficiently close in space at that time. In order to geometrically represent and investigate particle dynamics, the tools from complex networks are here employed. In particular, nodes correspond to levels, l_i, i.e. spanwise-groups of particles initially at the same y^+, thus resulting into $\mathcal{N}_v = N_l = 100$ nodes. By doing so, information of particles starting at the same wall-normal positions is enclosed in each node. Therefore, the metrics extracted from the network increase their statistical significance because they do not represent the dynamics of a single particle. For any time, a weight, $\mathcal{W}_{i,j}$, is associated to each link between a pair of nodes $\{i, j\}$, which takes into account the total number of connections shared by particles belonging to levels l_i and l_j. It is

worth to emphasize that we here explicitly employ the terms *connection* and *link* to indicate a relation between particles and nodes (that are groups of particles), respectively. An example of a time-varying network is shown in Figure 7.3(b), in which the particle dynamics represented in Figure 7.3(a) is geometrized into time-evolving networks (where nodes correspond to levels and the link thickness is proportional to link weight). As a result, particle dynamics is formally modelled by means of a time-varying weighted network, that is a sequence of $T^+/\Delta t^+$ weight matrices, defined as

$$\mathcal{W}_{i,j}(t^+) = \sum_{g \in l_i} \sum_{q \in l_j} \mathbb{I}_{g,q}(a_\chi, t^+) \mathcal{F}_{g,q}\,(d_{g,q}), \tag{7.1}$$

where $d_{g,q}$ is the modulus of the Euclidean distance between particles g and q, $i, j = 1, ..., \mathcal{N}_v$ correspond to node indices (with $\mathcal{W}_{i,j} = \mathcal{W}_{j,i}$), and $\chi = \{x, y, z\}$. In Eq. (7.1), the binary indicator function, $\mathbb{I}_{g,q}$, is equal to 1 if a particle g lies inside the ellipsoid of a particle q (or *vice versa*) at time t^+, and 0 otherwise. Accordingly, $\mathbb{I}_{g,q}$ explicitly depends on the ellipsoid size and time. Moreover, the window function, $\mathcal{F}_{g,q}$, is a weighting function taking into account the interaction strength between particle pairs $\{g, q\}$. In this work, $\mathcal{F}_{g,q}$ depends on $d_{g,q}$, namely the interaction strength between particles g and q is quantified via the pairwise Euclidean distance between g and q; however, other similarity or distance functions can be generally adopted.

Weighting particle interactions. As discussed for the characterization of the ellipsoid size, a_χ, the choice of $\mathcal{F}_{g,q}$ is conditioned to the specific problem under study. In this work, two configurations of window function $\mathcal{F}_{g,q}$ are explored.
First, in order to show the main features of the time-varying networks in a simple case, we set $\mathcal{F}_{g,q} \equiv \mathcal{F}^u = 1$, $\forall\{g, q\}$, where superscript u indicates a uniform window function; in other words, particle interaction is unweighted. In this way, each connection is equally weighted and the connection criterion is only driven by the ellipsoid size, $\{a_x, a_y, a_z\}$ – i.e., the main criterion is to check if particles come sufficiently close or not, at any time. As a result, for $\mathcal{F}_{g,q} = 1$, each link-weight exactly counts the total number of connections shared by particles belonging to pairs of nodes. To grasp the concept, Figure 7.4(a) shows a particle q (with coordinates X_q, Y_q and Z_q) and a 2D section of its ellipsoid (i.e., the ellipse) for $z = Z_q$, as well as the value of $\mathcal{F}^u = 1$ for each point inside the ellipse.
In the second configuration, we assume the intensity of the interaction between two particles decreases inversely proportional to the particle distance as

$$\mathcal{F}_{g,q} \equiv \mathcal{F}^w_{g,q} = \left(\frac{1}{d_{g,q}+1} - \frac{1}{d^{max}_{g,q}+1} \right) \Bigg/ \left(1 - \frac{1}{d^{max}_{g,q}+1} \right), \tag{7.2}$$

where $d_{g,q}$ is the Euclidean distance between particles g and q, given that g and q are connected (i.e., $\mathbb{I}_{g,q} = 1$, see Eq. (7.1)). If the ellipsoid is centred on a particle

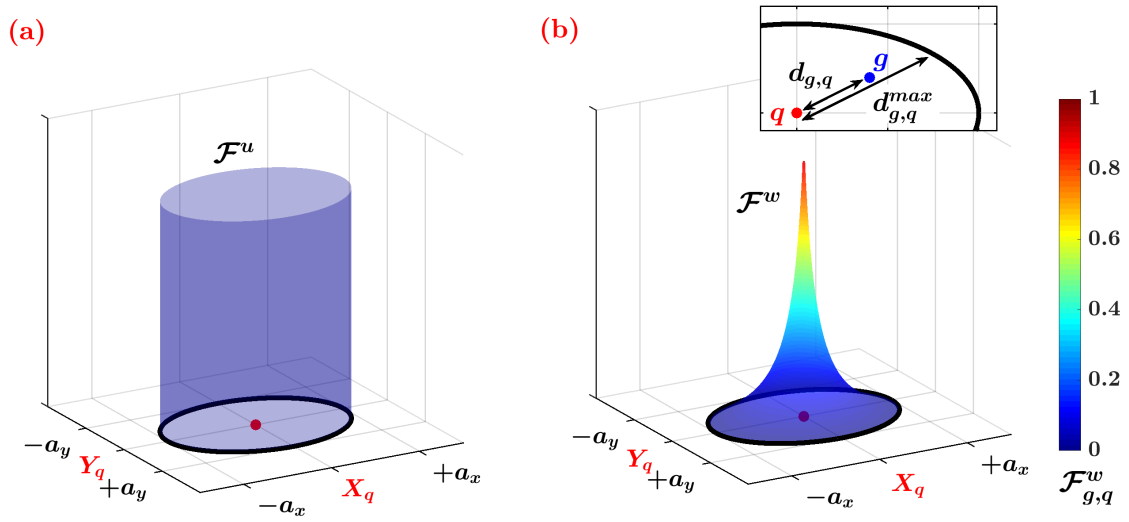


Figure 7.4: 2D section of the ellipsoid of a particle q (depicted as a red dot), with coordinates X_q, Y_q and Z_q, and the values of (a) $\mathcal{F} \equiv \mathcal{F}^u$ and (b) $\mathcal{F} \equiv \mathcal{F}^w$, for each point inside the ellipse obtained for $z = Z_q$.

q, $d^{max}_{g,q}$ is the distance from q and the point of intersection between the ellipsoid border and the straight line between g and q (the same criterion holds if the ellipsoid is centred on the particle g). Figure 7.4(b) shows the values of the function $\mathcal{F}^w_{g,q}$ for the ellipsoid centred in q, while the inset highlights the difference between $d_{g,q}$ and $d^{max}_{g,q}$. In this way, the window function is bounded as $\mathcal{F}^w \in [0,1]$, where the ellipsoid border represent an iso-value $\mathcal{F}^w = 0$, while $\mathcal{F}^w = 1$ is obtained if particle positions coincide (see Figure 7.4(b)). In this case, the anisotropy of the flow is explicitly considered as $d^{max}_{g,q}$ depends on the length of the ellipsoid semi-axes. By using a monotonically decreasing function for $\mathcal{F}_{g,q}$, the smallest spatial scales are more weighted than the largest ones, corresponding to the assumption that the smallest turbulent scales play a more significant role in particle mixing.

Finally, it is worth to mention that, since the network structure depends on particle relative positions, the accuracy of the interpolation method – employed to evaluate particle position from the velocity field (see Appendix B.5) – might play a role in the network statistics. In this work, a tri-linear interpolation is used as it was found to be sufficiently accurate to provide reliable Lagrangian statistics [65, 191]. Since connections are established based on the proximity criterion in which the size of the ellipsoid is proportionally adapted to the average distance between particles, small variations in the exact particle positions are supposed to barely affect the overall network structure. In the peculiar case in which a group of particles occupy the same position, the ellipsoid-based approach will capture the particle mutual proximity, either if particle distance are weighted (i.e., via $\mathcal{F}^w$) or uniform (i.e., via $\mathcal{F}^u$). In particular, the introduction of a non-uniform weighting function as

defined in Eq. (7.2) can then shed light on the presence of such peculiar cases, as the coincidence of particle positions correspond to $\mathcal{F}^w = 1$ (see also Figure 7.4(b)).

To summarize, the weight matrices, $\mathcal{W}_{i,j}(t^+)$, capture and inherit the pairwise particle positions at any time, t^+, by also retaining the information of the initially y^+ value of each particle. Specifically, a pair of nodes that is not linked at a given time (i.e., $\mathcal{W}_{i,j} = 0$) consists of two levels whose particles are not sufficiently close in space (namely, a link is absent in the network topology). On the other hand, non-zero $\mathcal{W}_{i,j}$ values quantify the intensity of the link between levels, namely the extent to which particles are close in space. By doing so, nodes do not vary with time (i.e., they still represent the same levels), while link weights depend on time. It should be emphasized that the links in the time-varying network are weighted either if the window function is uniform (i.e., $\mathcal{F} \equiv \mathcal{F}^u = 1$) or not (i.e., $\mathcal{F} \equiv \mathcal{F}^w$ as in Eq. (7.2)). In fact, the link weight takes into account the aggregated information of connections between all particles in each level, either if particle interaction is quantified via a uniform or non-uniform $\mathcal{F}_{g,q}$ function.

7.3 Lagrangian mixing analysis in a turbulent channel flow

In this section, the results of the time-varying weighted network approach to investigate turbulent mixing are reported for the two window function configurations, namely $\mathcal{F} \equiv \mathcal{F}^u = 1$ (see Section 7.3.1) and $\mathcal{F} \equiv \mathcal{F}^w$ (see Section 7.3.2). Throughout the results, the relative effect of the mean flow advection and the vertical mixing on particle dynamics is highlighted through the investigation of the network features.

7.3.1 Temporal characterization of turbulent mixing

The structure of the time-varying weighted network, $\mathcal{W}_{i,j}(t^+)$, is focused first. Figure 7.5 shows the weight matrices at six characteristic times and their corresponding network topology (for further visualizations see *Movie 1* in the Supplementary materials in Ref. [187]). The particle positions displayed in Figure 7.1(a)-(d) corresponds to the networks shown in Figure 7.5(a),(b),(d) and (f), respectively.

Since particles are initially arranged in a uniform grid at $x^+ = 0$ (see Figure 7.1(a)), each level is simply connected to levels close-in-space in the (y^+, z^+) plane at $t^+ = 0$, resulting in the diagonal weight matrix of Figure 7.5(a). For small times (see Figure 7.1(b)-(c)), the particle dynamics is led by an almost pure advective motion and the particle swarm takes the shape of a bow-like surface (at the very beginning, this surface reproduces the mean velocity profile, $U(y^+)$, in each (x^+, y^+) plane). Accordingly, the weight matrix exhibits a predominant diagonal

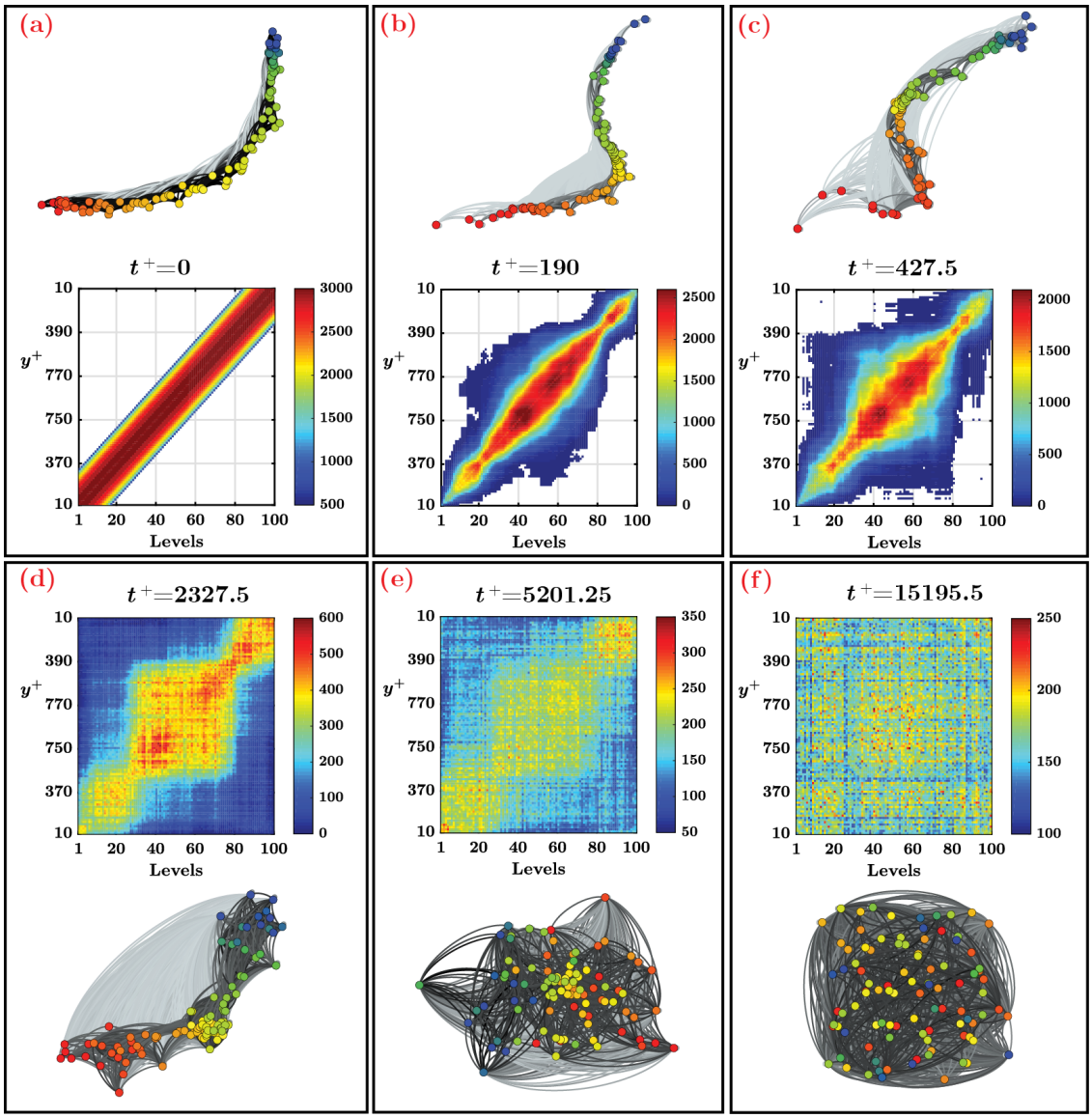


Figure 7.5: Time-varying weighted networks represented as weight matrices, $\mathcal{W}_{i,j}$, and their corresponding network topology for six representative times. The labels of the matrix ordinates indicate the y^+ value (relative to the closest wall, i.e. $y^+ \in [0,950]$) of the corresponding level reported in the matrix abscissas, while colors represent the weight of the links, $\mathcal{W}_{i,j}$. Network visualizations are obtained through the *OpenOrd* layout algorithm [130]; node colors indicate different y^+ values and range from red (i.e., level 1) to blue (i.e., level 100), while link weights are shown in a grey-scale, where strong links are in black and weak links are in light-grey.

pattern, and the network shows a tree-like elongated structure in analogy with the diagonal pattern of the corresponding $\mathcal{W}_{i,j}$, as shown in Figure 7.5(b). However,

wall-normal turbulent mixing enables out-of-diagonal links, and an increasing number of weak $\mathcal{W}_{i,j}$ values between initially distant levels appears in time (e.g., levels l_{10} and l_{75} in Figure 7.5(c)).
Sufficiently far enough downstream from $x^+ = 0$, all levels are interconnected with each other (e.g., see Figure 7.5(d)), due to the progressively enhanced transversal mixing. Nevertheless, the mean flow advection is still dominant over mixing at this stage: the initial linear diagonal structure of $\mathcal{W}_{i,j}$ evolves into a three-square diagonal pattern, where the central square is bounded between $y^+ \approx 500$ and $y^+ \approx 1400$ (namely levels l_{27} and l_{74}). This wall-normal coordinate corresponds to the y^+ value at which the mean shear, $\partial U/\partial y$, sharply decreases towards zero (see Figure B.2 in Appendix B.5). In fact, while particles initially located far from the walls experience an almost zero mean shear (thus moving downstream at a high mean velocity), particles close to the wall tend to form two long tails due to the large mean shear close to the wall (see Figure 7.1(c)). The advection process makes the tails progressively stretched along the walls as time increases, highlighting the effect of the mean shear on particle swarm. Therefore, the three-square pattern emerges as a consequence of the mean shear on the particle dynamics. From the point of view of the network topology, turbulence mixing – by enabling links between distant levels – has the effect to induce a clustered topology, namely a network geometry in which nodes tend to aggregate with each other. In fact, as illustrated in Figure 7.5(d), the nodes of the network at $t^+ = 2327.5$ tend to group based on their wall-normal coordinate, equivalently to the three-square pattern of $\mathcal{W}_{i,j}$.
Finally, at some time long after, turbulent mixing becomes as effective as streamwise advection. This is first manifested as a smoothing in the three-square pattern of $\mathcal{W}_{i,j}$ (Figure 7.5(e)), and later as a random-like structure (Figure 7.5(f)). This final state represents the Taylor dispersion regime (see Figure 7.1(d)), in which the streamwise particle distribution approaches a Gaussian distribution (see also *Movie 2* in the Supplementary materials in Ref. [187]). In this asymptotic stage, the network topology develops towards a strongly aggregated pattern with a random-like layout.

In order to highlight the richness of the information contained in the weighted network as time evolves, a scalar metric, $\mathcal{E}_p \equiv \sum_i \sum_j \mathcal{W}_{i,j}/2$, is introduced that is equal to the total number of particle connections established at each time. Figure 7.6 shows the behaviour of $\mathcal{E}_p$ as a function of t^+, in which it is possible to discern three temporal regimes:

I) The first regime ranges in $t^+ \in \left(0, T_{\mathrm{I}}^+\right]$, where $T_{\mathrm{I}}^+ \approx 450$ is the time-scale in which particle dynamics is primarily led by streamwise advection. Therefore, in this first regime particles are arranged in a bow-like shape, as shown in Figure 7.1(b). Since the wall-normal mixing is the main factor responsible for the activation of distant inter-levels links, in the first regime the total number of links is almost unchanged and only particles initially close in space

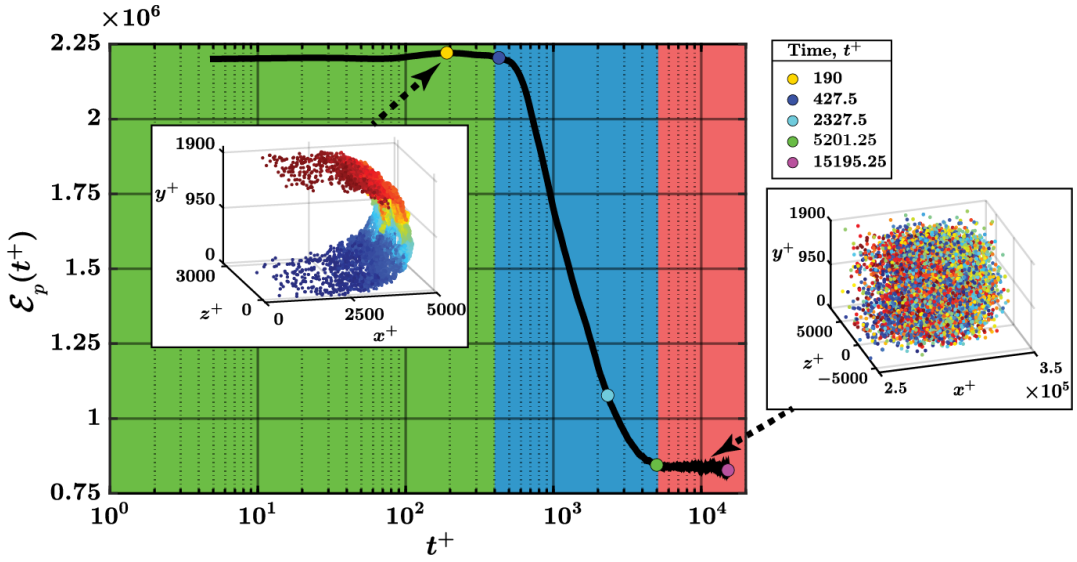


Figure 7.6: Total number of connections, $\mathcal{E}_p$, as a function of time, t^+. Three regimes are highlighted and distinguished by $T_{\mathrm{I}}^+ \approx 450$ and $T_{\mathrm{III}}^+ \sim 5200$. The two insets show the particle swarms in the first and third regimes, where particle colors indicate their starting level (as in Figure 7.1(a)). The values of $\mathcal{E}_p$ at the six times reported in Figure 7.5 are also illustrated as coloured circles.

are connected with each other (see also Figure 7.5(b)-(c)).

II) As time increases, the effect of mixing on particle dynamics progressively strengthens and $\mathcal{E}_p(t^+)$ decreases up to $t^+ \approx 5200$. In fact, if two particles are connected (i.e., each particle lies inside the reference ellipsoid of the other one), wall-normal mixing tends to move the two particles apart in the wall-normal direction. Accordingly, particle arrangement tends to deviate from the bow-like profile (typical of the first regime), and particles (dispersing in the y^+ direction) come across a region in which the particles are less dense (e.g., see Figure 7.1(c)). As a result, in this second regime, $\mathcal{E}_p(t^+)$ decreases as turbulent mixing progressively intensifies its effect on particle dynamics. In other terms, the second regime is an intermediate stage between an advection-dominant and a mixing-dominant regime.

III) At some time long after, the advection process and the transversal mixing are balanced and the particle dynamics approaches the Taylor asymptotic state. Hence, the third regime is characterized by a nearly constant value of $\mathcal{E}_p(t^+)$, because the particle swarm spreads in the streamwise and spanwise directions without showing any spatial pattern (see Figure 7.1(d)). According to our analysis, this third regime starts at time $T_{\mathrm{III}}^+ \approx 5200$, corresponding to the time-scale from which Taylor's dispersion analysis can be applied [189]. By

comparing T^+_{III} and the value of the integral wall-normal (Eulerian) mixing time-scale, which is approximatively equal to $14200\nu/u_\tau^2$ (see Section 7.2), we found that Taylor's analysis holds for $T^+_{\mathrm{III}}/14200 \gtrsim 0.37$ which is in excellent agreement with the value 0.4 reported in literature [188].

It is worth noting that, while the values of the integral wall-normal mixing time-scale and T^+_{III} can be estimated from a dimensional analysis of the conservation equation of a tracer, the value of T^+_{I} (i.e., the end of the first regime) is non-trivial. Indeed, T^+_{I} is governed by a transient dynamics in which the magnitude of the advection and mixing terms is not easy to quantify. By means of particle geometrization into networks, however, it is possible to easily distinguish the onset of both transient (i.e., T^+_{I}) and long-term regimes (i.e., T^+_{III}). It should also be noted that the chosen time resolution ($\Delta t^+ = 4.75$) is sufficiently accurate for the present analysis, as the network features smoothly evolve over time (see Figure 7.6).

To summarize, the network structures are directly affected over time by wall-normal turbulent mixing or, in general, by the interplay between mixing and advection. In particular, the effect of turbulent mixing on particle dynamics is captured by the total number of connections, $\mathcal{E}_p$, which is able to reveal to which extent wall-normal mixing breaks the initial particle arrangement towards the Taylor asymptotic state.

Further insights on wall-normal turbulent mixing

The total number of connections, $\mathcal{E}_p(t^+)$, is a useful and simple metric to effectively quantify the relative intensity of turbulent mixing and advection on particle dynamics. Indeed, the investigation of $\mathcal{E}_p$ provides concise insights into the ensemble behaviour of all levels at each time, and it enables to distinguish different advection-mixing regimes. Nevertheless, from the weight matrices $\mathcal{W}_{i,j}$ it is possible to extract much more detailed information. In fact – as shown in Figure 7.5(b)-(c) where low $\mathcal{W}_{i,j}$ values appear out of the main diagonal – the key effect of mixing is to promote the activation of links between nodes corresponding to distant levels (e.g., l_1 and l_{100}). Since particle geometrization into the network framework is based on the spatial proximity, the appearance of a link between two distant levels represents a peculiar event, which is important information, for instance when particles are involved in chemical reactions.

In order to characterize the appearance of such peculiar events, in Figure 7.7(a) we show the temporal behaviour of $\mathcal{W}_{i,j}(t^+)$ for six representative pairs of nodes. The node pairs selected involves levels close to the two walls (i.e., l_1 and l_{100}), the level close to the channel center (i.e., l_{50}) as well as an intermediate level, l_{25}. If we focus on how connections between particles in the same level change over time (that is the main diagonal of $\mathcal{W}_{i,j}$, with $i = j$), link activation starts as expected at the initial time (both for levels close to the wall, l_1, and at the center, l_{50}) and the

weight decreases towards an asymptotic average value. For the level pair {50,25} – namely particles initially started at $y^+ \approx 940$ and $y^+ \approx 465$, respectively – link activation starts quickly in time, because particles belonging to l_{25} and l_{50} do not experience a strong velocity gradient and turbulent mixing enables their connection after a short time interval. A link between the pair {1,25}, instead, appears only in the second regime at $t^+ \approx 570$, because during the first regime particles in l_1 experience a higher mean shear than particles in l_{25}, thus only when the mixing is strong enough a link appears between them.
Unexpectedly, as shown in Figure 7.7(a), a link between the two furthest levels, {1,100}, appears before a link between levels {1,50}. This can be explained by recalling that if particles in the bow-like swarm are moved apart due to wall-normal mixing, they come across a less dense region of other particles. Particles in l_1 are hence more likely to connect with particles in l_{100} than particles in level l_{50}, because when $\mathcal{W}_{1,100} > 0$ there is not spatial proximity between particles in l_1 and l_{50} as particles in l_{50} are mainly located far downstream (so that $\mathcal{W}_{1,50} = 0$). Therefore, link activation strongly depends on the flow features, namely both mixing and advection, and it reveals non-trivial results. In particular, since each node represents a spanwise-row of 100 particles, the activation of a link between two nodes takes into account all the pairwise connections between all the 100 particles in each node. By doing so, each link captures and highlights – through its weight – ensemble information about the dynamics of the spanwise-row set of particles in each node. It is worth noting that, for simplicity, we only explored the inter-relation between pairs of nodes, namely 2−tuples, but the weighted networks comprise the information of all n−tuples of nodes.

Finally, we focused on how the presence of different levels inside each ellipsoid varies over time due to the effect of mixing. In fact, any particle belonging to a level l_i is connected – due to spatial proximity – with a set of other particles belonging to different levels l_j, at each time. This is illustrated in Figure 7.2 where the presence of different levels inside the ellipsoid of a reference particle is highlighted by different particle colors. To quantify such variability at any time, we evaluated the standard deviation, σ_{l_i}, of all the indices $j = 1, \ldots, \mathcal{N}_v$ of levels l_j found inside the ellipsoids of all particles belonging to a reference level, l_i. In other terms, we quantify the variability of the presence of particles belonging to levels l_j, inside the ellipsoid of particles in l_i. For example, in the sketch of Figure 7.2, a higher standard deviation value would be obtained for the reference particle in green at time t_2 than at time t_1, because at t_2 there is more variability of particle colors (indicating different initial levels). In this way, we are able to quantify the efficiency of mixing between different levels.
In Figure 7.7(b) we show σ_{l_i} as a function of time for five representative levels, l_i. Due to the progressive strengthening of turbulent mixing, the values of σ_{l_i} generally increase with time. More in detail, since in the first regime advection tends to move particles apart in the streamwise direction, levels close to the walls

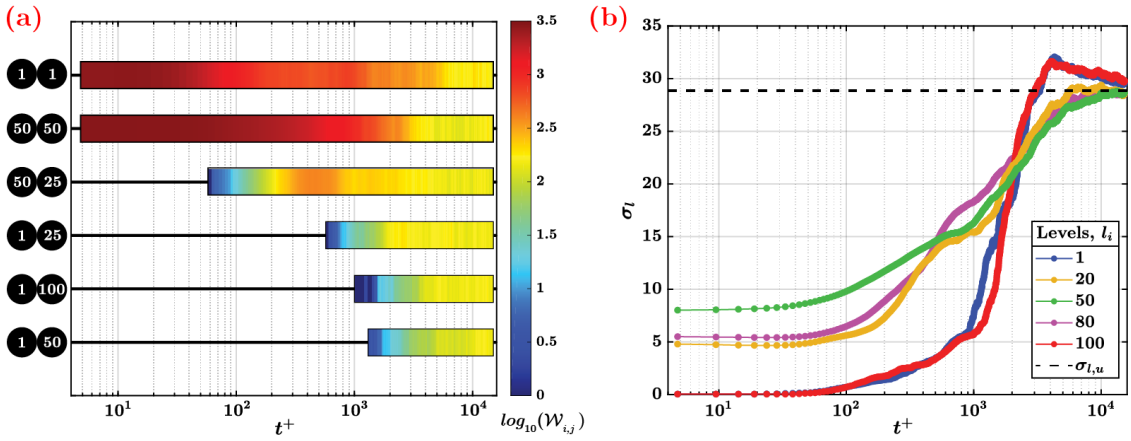


Figure 7.7: (a) Time evolution of $\mathcal{W}_{i,j}$ for six pairs of levels. Each horizontal bar corresponds to an entry of the $\mathcal{W}_{i,j}$ matrices, while link activation is highlighted by horizontal coloured bands, where color variations indicate the change of the link weight as $log_{10}(\mathcal{W}_{i,j})$. (b) Standard deviation, σ_{l_i}, of different levels inside the ellipsoid of particles in five representative levels. The horizontal dashed line indicates the standard deviation from a discrete uniform distribution in the interval $[1, \mathcal{N}_v]$, namely $\sigma_{l,u} \approx 28.87$.

display lower values of σ_{l_i} because they are unlikely to connect with other levels at higher y^+ (see Figure 7.1(b)). For long times, instead, all the σ_{l_i} approach the value of a uniform distribution, $\sigma_{l,u}$, because of strong turbulent mixing. However, only levels close to the wall (i.e., l_1 and l_{100}) show σ_{l_i} values larger than $\sigma_{l,u}$, as a consequence of their preferential linking with very distant levels (as previously discussed). It should be noted that the value of $\sigma_{l,u}$ can be analytically computed as the standard deviation of a uniform (discrete) distribution between 1 and $\mathcal{N}_v = 100$, that is $\sqrt{(\mathcal{N}_v^2 - 1)/12} \approx 28.87$. Moreover, pairs of levels at a similar distance from the wall (i.e., pairs $1 - 100$ and $20 - 80$) show analogous behaviour, since they experience similar dynamics.

Network sensitivity on the ellipsoid size

To conclude this section, a sensitivity analysis of the network features is performed for different size of the ellipsoid. Due to its ability to easily capture and summarize important features of the time-varying weighted network, the total number of connections, $\mathcal{E}_p$, is focused. For a fixed number of nodes, $\mathcal{N}_v$, the behaviour of $\mathcal{E}_p(t^+)$ – and in turn of the weighted network – basically depends on two modelling parameters: the number of particles in each level, and the constant of proportionality, α. Besides, the DNS spatial resolution could affect the network structure, since a coarse spatial resolution implies an inaccurate velocity field and, in turn, a poor

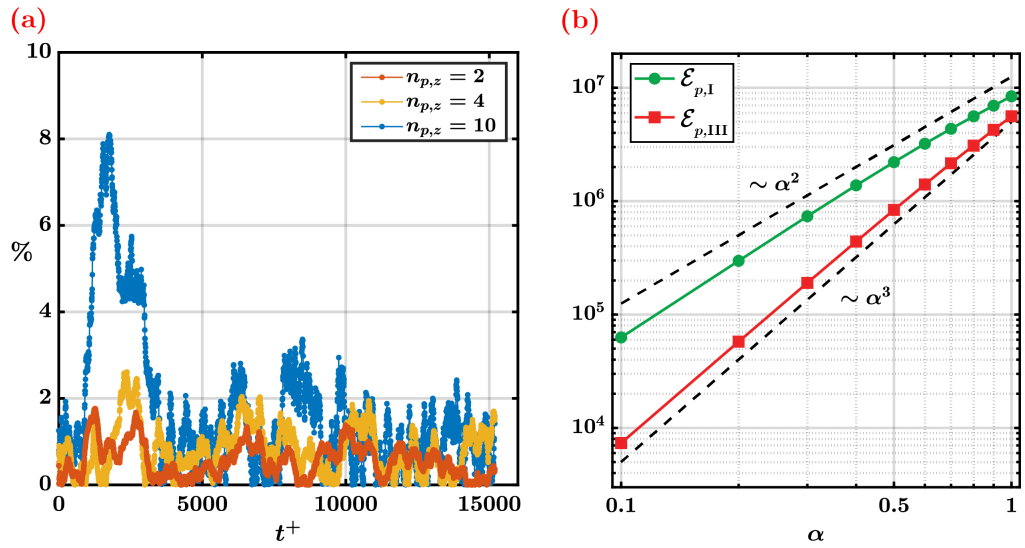


Figure 7.8: (a) Relative error (in percentage), $|1-n_{p,z}^2 \cdot (\mathcal{E}_p)_{n_{p,z}>1}/(\mathcal{E}_p)_{n_{p,z}=1}|$, between $\mathcal{E}_p$ for $n_{p,z} = 1$ (reference case) and $n_{p,z}^2 \mathcal{E}_p$ for $n_{p,z} > 1$. (b) Effect of α on $\mathcal{E}_p(t^+)$, where $\mathcal{E}_{p,\mathrm{I}}$ and $\mathcal{E}_{p,\mathrm{III}}$ are the average value of $\mathcal{E}_p(t^+)$ over $t^+ < T_{\mathrm{I}}^+$ and $t^+ > T_{\mathrm{III}}^+$, respectively. Dashed lines indicate the scaling as α^2 and α^3.

particle position resolution. However, the adopted spatial resolution is sufficient to provide a reliable particle position evaluation [65, 191].

To assess the robustness of the proposed approach, a parametric analysis on the number of particles in each level is first performed by keeping $\alpha = 0.5$ constant. In the reference case, each level comprises 100 particles, and by decreasing this value (keeping the number of nodes $\mathcal{N}_v = 100$ constant), the total number of particles released in the channel is reduced, thus weakening the statistical significance of the results. Here we decrease the number of particles initially released in the domain along, z, by a factor $n_{p,z} = \{2, 4, 10\}$, corresponding to a total number of released particles equal to $10^4/n_{p,z}$. By considering a fraction $1/n_{p,z}$ of the total number of particles, the decrease in the total number of connections $\mathcal{E}_p(t^+)$ is proportional to $1/n_{p,z}^2$. Figure 7.8(a) shows the relative error (in absolute value and in percentage) between $\mathcal{E}_p$ for $n_{p,z} = 1$ (reference case) and $n_{p,z}^2 \mathcal{E}_p$ for $n_{p,z} > 1$. As expected, the curve of $\mathcal{E}_p(t^+)$ scales down from the reference case proportionally to $1/n_{p,z}^2$, with a relative error below 10%. In fact, a decrease in the number of particles in each node implies a reduction of the total number of possible connections, which is proportional to the square of the number of particles (e.g., see Eq. (3.4)).

The effect of different α values, instead, is to vary the size of the reference volume for the activation of connections between particles. Hence, a higher (lower) value of α increases (decreases) the possibility that particles connect with each other. By keeping $n_{p,z} = 1$ and varying α in the range (0,1], the curve of $\mathcal{E}_p(t^+)$ is scaled but the values of T_{I}^+ and T_{III}^+ do not change. However, $\mathcal{E}_p(t^+)$ does not scale

as α^3 at any time, as one would expect by a 3D stretching of the ellipsoid, but the scaling depends on the regime. In fact, as shown in Figure 7.8(b), the mean value of $\mathcal{E}_p(t^+)$ for $t^+ < T_{\mathrm{I}}^+$ (namely $\mathcal{E}_{p,\mathrm{I}}$) scales as α^2, because in the first regime only a fraction of each ellipsoid is occupied by particles, resulting from the intersection of a bow-like surface with an ellipsoidal volume. Differently, the mean value of $\mathcal{E}_p(t^+)$ for $t^+ > T_{\mathrm{III}}^+$ (i.e., $\mathcal{E}_{p,\mathrm{III}}$) scales as α^3, because in the third regime particles are spread in all directions and particles occupy the entire volume of each ellipsoid (in a uniform way, as shown in Figure 7.7(b)).

7.3.2 Particle distance-weighted network analysis

The time-varying network built by equally weighting each particle connection (i.e., $\mathcal{F} \equiv \mathcal{F}^u = 1$) shows several remarkable features of the turbulent dispersion over time, as well as its effect on particle dynamics. However, in order to explicitly account for the pairwise distance between particles, the results for a non-uniform weighting function, $\mathcal{F}^w$, are here reported. We recall that $\mathcal{F}^w_{g,q}$ is a monotonically decreasing function of the pairwise distance $d_{g,q}$, as defined in Eq. (7.2) and illustrated in Figure 7.4(b). The ellipsoid is still used as a spatial proximity limit for particle connections (so that particles outside each ellipsoid are not considered since $\mathbb{I}_{g,q} = 0$), but now particles at smaller spacing inside each ellipsoid are more weighted (see Figure 7.4(b)).

Figure 7.9 shows the weight matrices, $\mathcal{W}_{i,j}$, for the $\mathcal{F}^w$ case at the same six representative times of Figure 7.5. As it emerges from the comparison of Figure 7.9 and Figure 7.5, the binary structure of networks for $\mathcal{F}^w$ at any time is the same as for $\mathcal{F}^u$. In fact, the diagonal and three-square patterns are also found for $\mathcal{F}^w$ in the first and second regime, respectively (see Figure 7.9(b)-(d)), while for large times patterns do not emerge (see Figure 7.9(f)). However, by using a non-uniform window function, the values of the link weights (i.e., the link colors in Figure 7.9) for the $\mathcal{F}^w$ case tend to be more intense along the diagonal, as shortest connections are weighted more.

In order to further characterize the time-varying weighted networks in both configurations, $\mathcal{F}^u$ and $\mathcal{F}^w$, four network metrics are investigated. The four metrics are selected in order to progressively highlight the main network features, ranging from a local (i.e., single nodes) to a global (i.e., the entire network) point of view: (i) the node centrality by evaluating the average strength, $\overline{\mathcal{S}}$; (ii) the relation between node pairs by evaluating the assortativity coefficient, r; (iii) the relation between node triples by evaluating the average clustering coefficient, $\overline{\mathcal{C}}$; and (iv) the relation between node $n-$tuples (of higher order) by computing all the shortest path length via the average path length, $\mathcal{L}$. The behaviour of these four metrics is shown in Figure 7.10 as a function of time, for both window function configurations, $\mathcal{F}^u$ and $\mathcal{F}^w$.

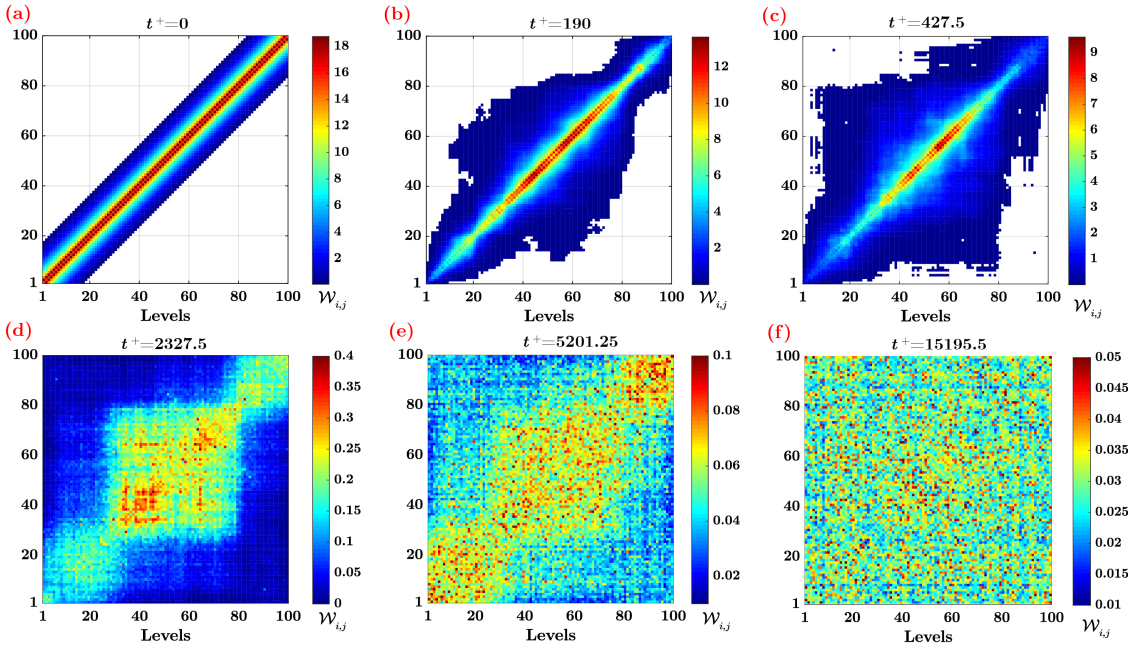


Figure 7.9: Time-varying weighted networks represented as weight matrices, $\mathcal{W}_{i,j}$, for a non-uniform window function, $\mathcal{F}^w$, at six representative times. Colours represent the weight of the links, $\mathcal{W}_{i,j}$ (for comparison with a uniform window function see Figure 7.5).

Average node strength. Figure 7.10(a) shows the average strength, $\overline{\mathcal{S}}(t^+)$, of the networks as a function of time (overline indicates average over all nodes). The strength of a node i in a weighted network quantifies the intensity of the relation between a node i and all other nodes (see Eq.(3.5)). In the case $\mathcal{F}^u = 1$, $\overline{\mathcal{S}}$ is related to the number of particle connections, $\mathcal{E}_p$, as $\overline{\mathcal{S}} = 2\mathcal{E}_p/\mathcal{N}_v$. The behaviour of $\overline{\mathcal{S}}$ as a function of t^+ for the two cases of $\mathcal{F}_{g,q}$ is essentially the same, because (as discussed in Section 7.3.1) the effect of turbulent dispersion on particle dynamics is fully captured by the network structure corresponding to $\mathcal{F}^u = 1$. However, the values of $\overline{\mathcal{S}}(t^+)$ for the case $\mathcal{F}^u$ are globally higher (about two order of magnitude) than the corresponding values for $\mathcal{F}^w$ (note that in Figure 7.10(a) there are two ordinate axes), as $\mathcal{F}^w$ ranges in the interval [0,1].
The difference in the temporal behaviour between the two $\mathcal{F}$ configurations is more evident in the first and the second regime, where $\overline{\mathcal{S}}$ for the $\mathcal{F}^w$ case (red curve in Figure 7.10(a)) decreases more rapidly. In the first regime, advection is dominant over mixing and most of the links are present between nodes initially close in space (i.e., along the diagonal of $W_{i,j}$). However, wall-normal turbulent mixing promotes the activation of links between initially distant nodes, as emerges from the appearance of out-of-diagonal links in Figure 7.5(b)-(c) and Figure 7.9(b)-(c), as well as in the increase of standard deviation in Figure 7.7(b). These out-of-diagonal

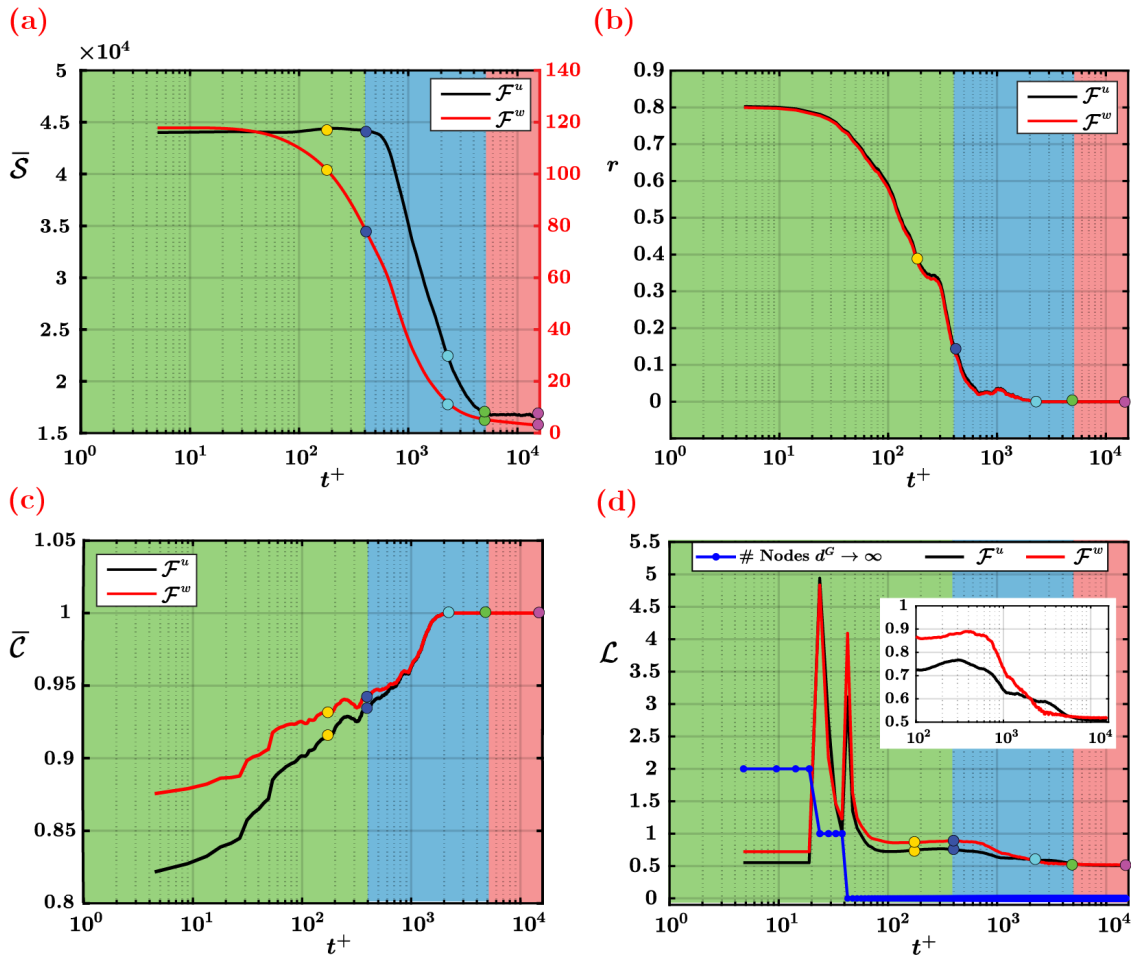


Figure 7.10: Network metrics as a function of time, t^+, for the two cases of window function, $\mathcal{F}^u$ (black lines) and $\mathcal{F}^w$ (red lines). (a) Average strength, $\overline{\mathcal{S}}$; the red ordinate axis on the right refers to the range of $\overline{\mathcal{S}}$ for the $\mathcal{F}^w$ case. (b) Weighted assortativity coefficient, r. (c) Average clustering coefficient, $\overline{\mathcal{C}}$. (d) Average path length, $\mathcal{L}$, and number of disconnected nodes in the networks (blue dotted line). The inset in panel (d) is a zoom of the decreasing behaviour of $\mathcal{L}$ with time. In all panels, coloured circles refer to the times reported in Figure 7.9, while background colors highlight the three advection-mixing regimes.

links correspond to connections between particles that are distant in space within an ellipsoid, thus corresponding to very low $\mathcal{F}^w_{g,q}$ values (see Figure 7.4(b)). In fact, if a particle g enters inside the ellipsoid centred in a particle q at a given time (with g and q belonging to distant initial levels), it is very likely that g is close to the border of the ellipsoid of q (where $\mathcal{F}^w_{g,q} = 0$). On the other hand, for the $\mathcal{F}^u$ case, connections between particles either close or distant in space are equally weighted. Therefore, the activation of distant links makes the values of $\overline{\mathcal{S}}$ decrease faster for

the $\mathcal{F}^w$ case than for the $\mathcal{F}^u$ case.

Assortativity coefficient. Following the concept of strength, in Figure 7.10(b), the assortativity coefficient, r, is shown as a function of t^+ (see Eq. (3.9)). We recall that a network is said to be assortative if $r > 0$ or disassortative if $r < 0$; if $r \approx 0$ the network is said to be non-assortative. In both $\mathcal{F}_{g,q}$ configurations, $r(t^+)$ similarly decreases from approximatively 0.8 at small times to zero at large times. In the first regime (especially at short times) nodes are primarily linked with other similar nodes, which implies that the networks are assortative; on the other hand, for long times, the networks lose any pattern, thus showing a non-assortative behaviour.

Clustering coefficient. The third metric analysed is the clustering coefficient, $\mathcal{C}$ (see Eq. (3.7)). Figure 7.10(c) shows the average clustering coefficient, $\overline{\mathcal{C}}(t^+)$, for the configurations $\mathcal{F}^u$ (black curve) and $\mathcal{F}^w$ (red curve). We recall that the clustering coefficient quantifies the probability that two randomly chosen neighbours of a node i (i.e., two nodes linked to i) are also neighbours, thus ranging in the interval [0,1]. In particular, a weighted clustering coefficient takes into account the interaction intensity between nodes comprising triplets [93]. As shown in Figure 7.10(c), $\overline{\mathcal{C}}$ increases to one for both configurations as time increases. In fact, for large times particles are well mixed with each other and it is very likely that nodes form triangles, namely nodes are locally very well inter-connected. The effect of weighting particle interactions via $\mathcal{F}^w$ is to enhance the local cohesiveness, especially in the first regime where most of the connections are between particles close in space. As a result, the values of $\overline{\mathcal{C}}$ are the highest in the $\mathcal{F}^w$ case, for which spatial proximity plays a more significant role.

Average path length. Finally, the average path length, $\mathcal{L}(t^+)$, namely the average of all shortest path lengths in the network (see Eq. (3.11)), is investigated. In general, a cost, $d^G_{i,j}$, is assigned to each link of the shortest path between node i and j. Here links are weighted according to the values of $\mathcal{W}_{i,j}$, which represents the intensity of spatial proximity between nodes, so that higher $\mathcal{W}_{i,j}$ implies closer distances. Accordingly, the shortest paths for the evaluation of $\mathcal{L}(t^+)$ are computed by using a weighting cost $\overline{\mathcal{W}}/\mathcal{W}_{i,j}$, where $\overline{\mathcal{W}}$ is the average link weight of the networks at any time [192]. The entries of the weighted matrix, $\mathcal{W}_{i,j}$, are normalized through $\overline{\mathcal{W}}$ in order to consider the change of $\mathcal{W}_{i,j}$ range as time increases (e.g., see the colorbar ranges in Figure 7.5 and Figure 7.9). Therefore, a high weight (i.e., a short distance between particles/nodes) means a low cost of the path.
Figure 7.10(d) shows the behaviour of $\mathcal{L}$ as a function of t^+. The networks during the first regime are able to display high $\mathcal{L}$ values, since two nodes are reachable via a high-cost path. By inspecting the values of $d^G_{i,j}$ we found that outliers in

$\mathcal{L}$ for $t^+ \in [20,50]$ are due to the nodes starting very close to the walls (namely, l_i with $i = \{1,100\}$). In the first regime, nodes starting very close to the walls are disconnected from the other nodes in the network (see blue-dotted curve in Figure 7.10(d)), namely $d^G_{1,j} = d^G_{100,j} = \infty$, because of a strong effect of mean shear on particle positions (see also Figure 7.5(b) and Figure 7.9(b)). Since disconnected nodes are conventionally excluded from the computation of the average path length [80], the peaks in $\mathcal{L}$ appear when nodes l_1 and l_{100} are linked to the other nodes in the network, i.e. at $t^+ \approx 25$ and $t^+ \approx 40$, respectively. From the point of view of the network topology, the re-attachment of the disconnected nodes implies high-cost shortest paths, as the values of $d^G_{i,j}$ for nodes $i = \{1,100\}$ are two orders of magnitude larger than $d^G_{i,j}$ for all the other nodes. As a consequence, very high values of $\mathcal{L}$ are detected. It should be noted that, although the effect of the mean shear is evident in the first regime, disconnected nodes are not expected for $t^+ \to 0$, as follows from the initial particle arrangement.
For long times, instead, $\mathcal{L}$ tends to decrease as an effect of turbulent mixing. By comparing the two configurations, $\mathcal{F}^u$ and $\mathcal{F}^w$, the values of $\mathcal{L}$ in the first two regimes are higher for the $\mathcal{F}^w$ case, because (as mentioned for the assortativity and the clustering coefficient) the window function highlights the local mixing by enhancing the spatial proximity. In the third regime, instead, both $\mathcal{F}_{g,q}$ configurations approach the same constant $\mathcal{L}$ asymptotic value (due to the link cost normalization).

To sum up, the analysis of the metrics at different network levels (from the single node, to node pairs and triples, as well as the shortest paths) of the time-varying weighted network, is able to unveil both the main features of particle dynamics and the presence of extreme cases. In particular, by exploiting a non-uniform function for particle interaction intensity, it is possible to highlight the effects on particle dispersion due to the different spatial proximity between particles.

7.4 Summary and future outlooks

The proposed Lagrangian network-based approach is exemplified by means of a DNS of a turbulent channel flow, where the dynamics of fluid particles is captured through a spatial proximity criterion. A time-varying weighted network is built by assigning to each node a spanwise-row of particles initially started at the same wall-normal coordinate, while link weight depends on the particle proximity at each time. The resulting time-varying weighted network is fully able to inherit the non-trivial time sequence of connections between (groups of) particles, which emerges due to turbulent motion. Indeed, we can identify in a straightforward way the characteristic regimes of particle dynamics, the appearance of peculiar events (e.g., the time of first contact between initially distant levels), as well as the intensity

of wall-normal turbulent mixing (quantified by the total number of connections). Accordingly, the potential of the Lagrangian-based networks is twofold, since the time-varying weighted network – represented by the weight matrices – captures in a unique framework both the qualitative spatial features of the particle swarm and the strength of turbulent mixing. Our approach is computationally affordable for a typical number of tracers of the order of $10^4 - 10^5$ (see Appendix A), and it is also suitable for experimental techniques such as particle tracking velocimetry.

Thank to complex network versatility, particle geometrical representation into time-varying networks can easily be extended to other setups and/or configurations. For instance, other turbulent flows can be investigated, such as flows in presence of thermal inhomogeneity, density stratification (i.e., buoyancy effects) or mass diffusivity (i.e., different Schmidt numbers). Besides passive scalars, other tracers as inertial particles can be employed in future works, thus also highlighting the feedback effect of particle inertia on turbulence. The effect of a different Reynolds number on the network features is another key issue that will be focused on future research. Indeed, the Reynolds number plays a key role to quantify the relative intensity of turbulent mixing and mean flow advection, which needs to be addressed.

Chapter 8

Conclusion and future perspectives

The current possibility to access a huge amount of numerical and experimental data of turbulent flows, have paved the way to the development of novel theories and methodologies. Complex networks, in particular, have revealed to be a powerful tool that is able to effectively take advantage from *big data*, by unravelling non-trivial issues in many research areas including fluid dynamics. Indeed, in the last two decades, complex networks have broadened the horizons of classical statistical analysis, by providing a different perspective – made up of interconnections between constitutive elements – on complex systems.

By exploiting the potential of complex networks theory, the spatio-temporal characterization of wall-bounded turbulence was focused as the main aim of the present Thesis. To put our aims in an interrogative sentence: how can wall turbulence characterization benefit from network science? And, which additional information do complex networks provide with respect to classical statistics? These issues were tackled by exploring three possible formats of raw-data, namely (i) one-point time-series, (ii) Eulerian spatio-temporal fields and (iii) Lagrangian particle trajectories. According to each data format, three approaches were carried out, namely a visibility graph analysis of time-series, a spatial network approach based on two-points correlation, and a time-varying network approach to particle dispersion. Fully developed turbulent channel flows and a passive scalar plume in a turbulent boundary layer are the two flows that were focused. In particular, numerically simulated data were mainly exploited in our analyses, but experimental measurements were also addressed. In fact, although they are limited to low or moderate Reynolds numbers, numerical simulations provide a unique instrument to gain high-resolution turbulence data both in space and time. For each of the three pursued approaches, substantial efforts were carried out to highlight turbulence features through the physical interpretation of the network structure and *ad hoc* selected network metrics.

A summary of the main contributions obtained in the Thesis is reported as follows.

Contributions to time-series analysis:

- For the first time, a visibility network-based analysis is carried out to characterize (i) the wall-normal structure of the velocity field in a turbulent channel flow, and (ii) the dispersion of a passive scalar plume in a turbulent boundary layer. By means of the network analysis, two important features of a time-series are highlighted: the occurrence and (relative) intensity of peaks and the presence of small fluctuations. In this way, higher-order information on the temporal structure of turbulence time-series are captured through the visibility algorithm. Indeed, concerning extreme events, we are able to straightforwardly discern between peaks and outliers and their temporal occurrence in the signals. Differently from higher-order statistics relying of the series PDF (e.g., the kurtosis), the visibility networks – and, in turn, the related metrics – strongly depend on the temporal structure of the signals.
- Particular attention is paid in this work to the physical interpretation of the network metrics with respect to the flow dynamics, so that specific trends of the network metrics are related to turbulent mechanisms involving the channel flow and the plume dynamics. With this aim, substantial efforts are made to relate the network features to the temporal structure of the series. This fundamental issue is often disregarded and minimized in the current literature, since the meaning of the metrics is often merely interpreted as a network feature rather than a signal feature.
- Limitations of the study are mainly related to the need of properly time-resolved time-series, either numerically of experimentally. From the point of view of the visibility algorithm, a computationally efficient ®MATLAB code is developed in this work to speed up the network construction (if a large datasets of time-series have to be mapped into networks). On other hand, the main advantages of the visibility algorithm rely on its extremely simple implementation as well as, different from other techniques, the lack of any *a priori* setting parameter.

Contributions from the Eulerian viewpoint:

- A spatial network analysis is proposed for the first time to spatially investigate a three-dimensional turbulent channel flow through a correlation-filtering approach. Although the idea to threshold the (Pearson) correlation coefficient (by retaining its highest values) to built spatial networks was borrowed by climate network analysis, here a 3D study of an inhomogeneous and anisotropic turbulent flow is newly introduced.

- In this work, the idea typical of climate analyses to assign the nodes to specific spatial locations and activate links via a threshold of the highest correlation coefficients is extended for the first time to a 3D turbulent flow setup. Accordingly, we introduce the idea that a network node corresponds to a finite volume in a spatial 3D domain. This conceptual extension is motivated by the combination of the flow inhomogeneity and the three-dimensional domain investigated.
- By focusing on the streamwise velocity field, the spatial network analysis reveals the presence of inter-wall and intra-wall teleconnections(in all directions) between spatially extended regions. The appearance of teleconnections between distant regions in the domain is one of the most significant results, since teleconnections emerge due to the ability of complex networks to retain the spatial information of high correlation coefficients. This information, instead, is typically hidden by the spatial averaging of the correlation coefficient values. Additionally, the most important nodes in the network are found to group in space and they result to be associated to the persistence of streamwise velocity streaks.
- The main limitations to this approach are due to the size of the resulting spatial network, since the number of nodes correspond to the number of spatial locations (i.e., grid points). The computational and storage cost is therefore the main drawback of the method (e.g., see Appendix A). To partially overcome this issue, different computational (and conceptual) strategies can be employed. Another important issue on the method concerns the temporal length of the signals used for evaluating the correlation coefficient. Due to the exploratory nature of our work, a time window that is higher than one channel flow through time is used. Since turbulent features – such as coherent motion – display their own temporal living scales, a more detailed sensitivity analysis on the total temporal window surely deserves further investigation.

Contributions from the Lagrangian viewpoint.

- A geometrical representation of Lagrangian particle dispersion in a turbulent channel flow is provided, for the first time, by exploiting a time-varying network approach. The novelty of our work with respect to the available literature is twofold. From one side, we introduce a time-varying network approach to geometrically represent particle dispersion, which highlights the transient phenomena of particle mixing and advection. In other network-based works, the temporal evolution of particle dynamics does not explicitly emerge, as links are evaluated in a time interval. On another side, for the first time we employ the network

framework in a three-dimensional, inhomogeneous and anisotropic turbulent flow. Most of the existing works, instead, have mainly focused on two-dimensional setups or surrogate flows of geophysical applications.

- To build our network-based framework, a well-established spatial proximity criterion is employed between particle pairs. The present formulation can be viewed as a generalization of the existing proximity-based network approaches and allows one to quantify the intensity of the particle interaction via a weighting function. This can be effectively used in many contexts (e.g., in chemical reactions) in which particles actively interact with each other.
- This work, therefore, contributes to provide a unique framework that is able to retain non-trivial information on particle dispersion. First, characteristic mixing-advection regimes can be easily identified through the network approach. Specifically, the identification of the transient time in which the advection-dominant regime ends, is a considerable outcome of our work. Moreover, by means of the Lagrangian network formulation, different metrics can be exploited to highlight the effects and intensity of turbulent mixing on particle dynamics.
- As for the previous network methodologies, the main limitation to this approach could be represented by the computational cost, since it is required to build a graph (with size equal to the number of particles) at each time. However, in this case, we show that a limited number of particles and a rather large time step are sufficient to capture the essential features of turbulent mixing with smooth temporal behaviours of the metric values. As a result, our network approach is sufficiently robust to allow a reduction of the computational cost, thus overcoming the main drawback of this formulation.

In conclusion, the complex networks show the advantage to highlight spatial and temporal features of turbulent flows that would be difficult to capture by means of other techniques. In fact, spatio-temporal insights into the turbulence mechanisms can be unravelled – differently from classical approaches – by focusing on the extent to which the constitutive elements that represent a turbulent flow (e.g., time-series or particle trajectories) are inter-related with each other. The resulting complex texture of interconnections is thus able to retain advanced information with respect to other classical statistics, e.g., the temporal structure of time-series that is lost by PDF-based statistics, as well as the directionality of (high) cross-correlation values that is typically lost through the average correlation coefficient.

Future outlooks. The aforementioned contributions, together with case by case limitations, represent different pieces of a broader picture of turbulent flow analysis via complex networks. Accordingly, the present Thesis provides one of the

first efforts to bring together two large research worlds – i.e., turbulence and network science – but several additional work deserves to be carried out and numerous achievements can then come forth. First, a thorough study to relate the network metrics and spatio-temporal aspects of the flow dynamics needs to be addressed, for all the three approaches. By providing an interpretation of the networks in the view of the main wall-normal turbulence features, this issue starts to be faced in this work, but still deserves further fluid dynamics insights. The physical interpretation of network metrics can then help in the identification and characterization of coherent motions and extreme events, which represent crucial topics in turbulence research. Additionally, a deeper understanding of the relation between spatio-temporal flow fields and the corresponding complex networks, can foster the development of novel low-order methods to reduce the computational costs (which can be particularly useful in spatial network analysis).

The ability of complex networks to retain local and global information of the mapped turbulent fields, can allow one to exploit network tools for flow modelling and control. In fact, thanks to the wide spectrum of metrics developed so far in network science, critical regions (e.g., high turbulence intensity) can be identified for flow control, especially in the Eulerian framework. Moreover, the network formulation can turn out to be a synthetic tool for turbulence modelling in different contexts, such as the dispersion of particles and persistent scaling laws. In view of this, the visibility and Lagrangian time-varying approaches could represent suitable frameworks for the development of network-based turbulence models.

An extension of the present work can also involve higher-order network formulations, namely the analysis of turbulent flows by means of multilayer networks. In fact, the multilayer approach can be pursued, e.g., if multivariate time-series are available or, in general, to highlight the cross-relation between networks built on different quantities (e.g., different velocity components).

Finally, although this Thesis focuses on wall turbulence, the network-based analysis can be extended to other flow configurations. For example, turbulent flows with thermal or density inhomogeneity can be focused for each of the three approaches, either through numerical simulations or experimental measurements (e.g., involving geophysical turbulence). Multiphase flows can also be addressed, especially in the Lagrangian network framework in which particles with different chemical-physical properties (e.g., particle inertia) can be employed to investigate turbulent mixing. Additionally, unsteady and transitional turbulence can be studied by means of complex networks. The spatio-temporal variations in the flow fields due to unsteady motion or a transition, indeed, could be captured by sudden changes in the network metrics, as it has been found out in other research areas, such as financial time-series.

In conclusion, due to the versatility and wide range of possible applications, the proposed network-based approaches can lead to relevant implications in several industrial and environmental fluid dynamics contexts.

Appendix A

Computational cost breakdown

In order to provide a quantitative overview of the computational costs associated to each of the three approaches carried out in this work, a report of the computational time to build and store the adjacency matrix is exemplified in Table A.1. The size in terms of number of nodes, $\mathcal{N}_v$, and time-frames employed, $T/\Delta t$, are selected to be the same as those in the main analyses.

Table A.1: Computational cost breakdown for different network approaches. The size, $\mathcal{N}_v$, and times, $T/\Delta t$, correspond to those pursued in the network analysis. Performances are evaluated for a machine with an Intel® Core™ i7-4790, CPU 3.60 GHz, 8 cores, 32 GB RAM, ®MATLAB R2018a.

	Visibility		**Eulerian**		**Lagrangian**	
$\mathcal{N}_v$	10^3	10^5	4×10^6		10^4	
$T/\Delta t$			5×10^3	10^4	3.2×10^3	3.2×10^2
Network building time (seconds)	4.5×10^{-3}	5.5×10^{-1}	7.5×10^5	1.5×10^6	6.8×10^3	6.8×10^2
Storage $\mathcal{A}_{i,j}$ (MB)	2.5×10^{-2}	8.5×10^{-1}	8×10^3	11×10^3	5×10^3 (2.5×10^2)	5.8×10^2 (8.1)

For the visibility approach, time-series extracted from a normal distribution (representative of the typical computational times of turbulence signals) are mapped into networks via the optimized code in Ref. [128], and times are averaged over 10^3 realizations. For the Eulerian approach, performances are evaluated for $\theta = 0.8$ and 0.7 for $T/\Delta t = 5000$ and 10000, respectively, and the computation times include the saving operation in the hard drive. Finally, for the Lagrangian case, performances are evaluated for $\alpha = 0.5$ (as in the main analysis). It should be

noted that the computational times reported in Table A.1 are obtained for a serial-computing approach. However, a distributed-computing approach (i.e., by running multiple routines concurrently) can lead to a significant reduction of computational costs. For instance, in the Eulerian case, times have been approximatively halved by running two concurrent ®MATLAB workspaces.

Appendix B

Description of turbulent channel flow DNSs

Some of the contents presented in this Appendix have been previously published. Below are provided the references:

G. Iacobello, S. Scarsoglio, J.G.M. Kuerten, and L. Ridolfi. "Spatial characterization of turbulent channel flow via complex networks". In: *Physical Review E* 98.1 (2018), p. 013107. DOI: 10.1103/PhysRevE.98.013107.

G. Iacobello, S. Scarsoglio, and L. Ridolfi. "Visibility graph analysis of wall turbulence time-series". In: *Physics Letters A* 382.1 (2018), pp. 1–11. DOI: 10.1016/j.physleta.2017.10.027.

G. Iacobello, S. Scarsoglio, J.G.M. Kuerten, and L. Ridolfi. "Lagrangian network analysis of turbulent mixing". In: *Journal of Fluid Mechanics* 865 (2019), pp. 546–562. DOI: 10.1017/jfm.2019.79.

B.1 Introduction

This Appendix contains details on the direct numerical simulations (DNSs) of turbulent channel flow at different Reynolds numbers, exploited for the computation of spatio-temporal fields used in the network analysis. Four different simulations are performed according to the frictional Reynolds number, Re_τ.

Table B.1 summarizes the main features of each dataset as used in the network analysis, as well as the relative chapter of the results in which the data are used. In each section of this Appendix, details on the numerical techniques as well as the spatio-temporal discretization are provided as a complementary material to the result sections.

Table B.1: Main parameters of the DNSs exploited for the network analysis. The first column refers to the Chapter in which the corresponding data are used.

Chpt	Re_τ	$L_{x,DNS}$	$L_{z,DNS}$	Δx^+	Δz^+	Δy^+_{max}	Δt^+_{DNS}	T^+
5	1000	$8\pi H$	$3\pi H$	12.26	6.13	6.16	0.325	1298.3
5-6	180	$4\pi H$	$4/3\pi H$	15.6	2.6	2.95	0.045	112.5 – 450
6	590	$8\pi H$	πH	19.31	4.83	7.24	0.07375	370
7	950	$2\pi H$	πH	7.8	3.9	7.8	0.095	15200

B.2 Details of the DNS at $Re_\tau = 1000$

In this Section we report the details of the direct numerical simulation (DNS) of a turbulent channel flow at $Re_\tau = 1000$. The data from this simulation are exploited to build visibility networks at different spatial positions from the velocity field (see the results in Chapter 5, Section 5.3).

The velocity field are extracted from a DNS of a fully developed turbulent channel flow, available from the Johns Hopkins Turbulence DataBase (JHTDB) [40, 193, 194]. The simulation is performed at the friction velocity Reynolds number $Re_\tau = Hu_\tau/\nu = 1000$, where $H = 1$ is the half-channel height, $\nu = 5 \cdot 10^{-5} U_b H$ is the fluid viscosity, $U_b = 1$ is the bulk channel velocity, and $u_\tau = 5{\cdot}10^{-2}$ is the friction velocity (all physical parameters are dimensionless). Periodic boundary conditions in the streamwise, x, and spanwise, z, directions are adopted, while the no-slip condition is imposed at the top and bottom walls. Once the statistically stationary conditions were reached, the simulation was carried on for approximately one flow-through time, $t \in [0{,}26]H/U_b$, with a storage temporal step $\delta t = 0.0065$. Thus 4000 temporal frames are available. Velocity and pressure fields were computed over the physical domain, $(L_x, L_y, L_z)_{DNS} = (8\pi H, 2H, 3\pi H)$, and stored with a grid resolution (in the physical space) equal to $(N_x, N_y, N_z)_{DNS} = (2048, 512, 1536)$. Further details on the simulation are available online (`https://doi.org/10.7281/T10K26QW`) as well as in Ref. [194].

In this work, in order to speed-up the computation, a subset of the domain was taken into account. In the wall-normal direction, y, due to the geometric symmetry, only grid-points from the bottom wall to the half-channel height are considered. In particular, a finer discretization is taken close to the wall while grid-points with gradually increasing spacing are taken towards the center of the channel. Differently, along the homogeneous directions, x and z, a coarse uniform discretization is adopted, which is sufficient to guarantee the statistical stationarity of the network results.

The selected grid-points are reported below (according to the labelling of the online database) in the form $(a : d : b)$, where a and b are respectively the first

and the last index of a uniformly spaced interval, and d is a grid step size (e.g., $(1:2:9)$ takes the indices $\{1;3;5;7;9\}$):

- streamwise direction, $X_i = (0:32:2016)$;
- wall-normal direction, $Y_i = \begin{cases} (0:1:21) \\ (23:2:39) \\ (42:3:54) \text{ and } (58:3:79) \\ (84:5:169) \\ (179:10:239) \\ 255, \text{ i.e. } y^+ = 996.3; \end{cases}$
- spanwise direction, $Z_i = (110:128:1518)$.

As a result, the selected sub-domain size is $(N_x, N_y, N_z) = (64, 70, 12)$, where the first grid-point $Y_i = 0$ corresponds to the wall coordinate $y^+ = 0$ (i.e, the wall).

B.3 Details of the DNS at $Re_\tau = 180$

In this Section we report the details of the direct numerical simulation (DNS) of a turbulent channel flow at $Re_\tau = 180$, which is exploited to build the spatial networks from the velocity field (see the results in Chapter 6).

The Navier-Stokes equations (continuity and momentum equations) are solved for incompressible flow,

$$\nabla \cdot \boldsymbol{u} = 0, \tag{B.1}$$

$$\frac{\partial \boldsymbol{u}}{\partial t} + \frac{1}{\rho}\nabla p_t = \boldsymbol{F} - \boldsymbol{\omega} \times \boldsymbol{u} + \nu \nabla^2 \boldsymbol{u}, \tag{B.2}$$

where $\boldsymbol{u} = (u, v, w)$ is the velocity of the fluid, $\boldsymbol{\omega} = \nabla \times \boldsymbol{u}$ is the vorticity, $p_t = p + \frac{1}{2}\rho \boldsymbol{u}^2$, ν and ρ are the kinematic viscosity and mass density of the fluid, p is the periodic part of the static pressure. Here, bold notation indicate vector quantities. The term $\boldsymbol{F}$ corresponds to the mean driving force per unit mass, which is uniform in space and in the streamwise direction, and chosen constant in time, in such a way that the Reynolds number based on the friction velocity is equal to 180.

The numerical approach is based on the method employed by *Kim et al.* [27], but with the time integration method by *Spalart et al.* [195]. A Fourier-Galerkin approach is used in the two periodic directions, while a Chebyshev-tau method is applied in the wall-normal direction. Instead of the velocity components, the wall-normal component of the vorticity vector and the Laplacian of the wall-normal velocity component are the dependent variables. In this way the incompressibility condition of Eq. (B.1) is automatically satisfied. The non-linear terms in the Navier-Stokes equation (B.2) are calculated in physical space by fast Fourier transform

(FFT) with application of the 3/2 rule in both periodic directions. A combination of a three-stage second-order accurate Runge-Kutta method and the implicit Crank-Nicolson method is chosen according to Ref. [195]. This method has been used and validated extensively at frictional Reynolds numbers ranging between 150 and 950 [65, 178, 191, 196, 197].

The time step, Δt, used in the simulation equals $2.5 \times 10^{-4} H/u_\tau$, which implies that $\Delta t^+ = \Delta t \cdot u_\tau^2/\nu = 0.045$ in wall-units. The number of time steps in the simulation is 5000, which corresponds to a time $Tu_\tau/H = 1.25$, or $T^+ = 225$ in wall-units. The DNS domain sizes, $(L_x, L_y, L_z)_{DNS}$, are equal to $4\pi H$, $2H$ and $\frac{4}{3}\pi H$ in the streamwise, wall-normal and spanwise directions, (x, y, z), respectively. The number of Fourier modes in the streamwise and spanwise directions equals 384 and 192, respectively, while 193 grid points in the wall-normal direction are exploited. This implies that in physical space the number of simulated grid points equals $(N_x, N_y, N_z)_{DNS} = (576 \times 193 \times 288)$. The DNS grid spacing (in wall-units) in the streamwise and spanwise directions are $\Delta x^+_{DNS} \approx 3.9$ and $\Delta z^+_{DNS} \approx 2.6$, respectively. In the wall-normal direction, the following spacing is taken

$$y_i = 1 - \cos\left[\frac{(i-1)\pi}{N_{y,DNS} - 1}\right], \tag{B.3}$$

where $i = 1, \ldots, N_{y,DNS}$ are the indices of the grid points in the wall-normal direction.

However, in order to have a manageable network size, the computational domain and the spatial discretization are reduced in the streamwise and spanwise directions. This choice is justified by the fact that the metrics analysed (see Chapter 6) are mainly dependent on the wall-normal coordinate (i.e., the inhomogeneous direction), so that this operation does not alter the significance of the results. Accordingly, with respect to the DNS domain size and discretization, one out of every four grid points in the x direction are selected, which results in 144 equally spaced grid points (instead of 576). In the z-direction, 150 consecutive grid points (instead of 288) with the same resolution of the DNS are considered. As a result, the streamwise spacing increases as $\Delta x^+ = 4\Delta x^+_{DNS}$, while the spanwise size of the domain reduces to $L_z = L_{z,DNS} \cdot N_z/N_{z,DNS} = 25/36\pi H$. In the wall-normal direction, instead, the grid points corresponding to the walls (i.e., $y^+ = 0$ and $y^+ = 360$) are excluded, since in those locations the velocity time-series are constantly zero, so that $N_y = 191$ (instead of 193). By doing so, the final domain considered for the network analysis is only periodic in the x direction. The total volume of the domain changes from $V_{tot,DNS} \approx 105.3H^3$ to $V_{tot} \approx 54.8H^3$, while the final spatial discretization is $(N_x, N_y, N_z) = (144, 191, 150)$.

It should be noted that the whole length of the domain is maintained in the streamwise and wall normal directions. In fact, they are the directions of advection and inhomogeneity of the flow, respectively, thus playing a crucial role in the channel turbulent dynamics. For instance, by retaining the whole streamwise length of

the domain, it is possible to entirely capture elongated turbulent structures such as streaks, which have scales of the order of L_x.

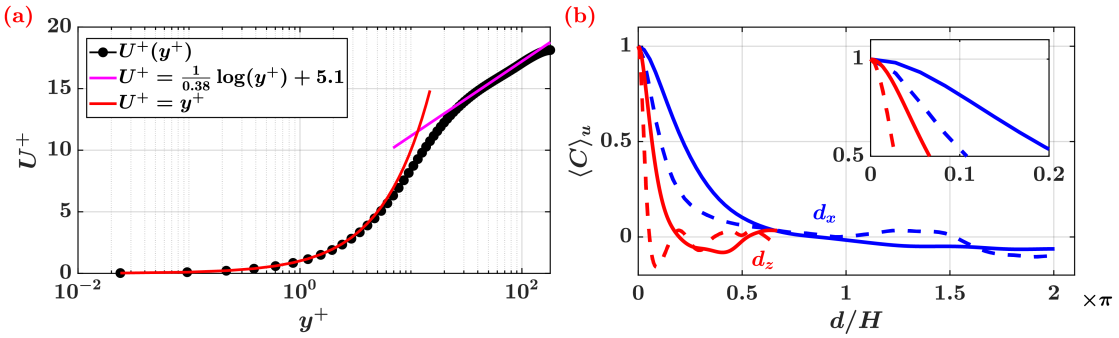


Figure B.1: (a) Mean velocity profile (in wall-units) as a function of the wall-normal coordinate, y^+. The law-of-the-wall is also shown. (b) Average two-points correlation, $\langle C \rangle_u$, of the streamwise velocity, u, at $y^+ = 3.5$ (solid lines) and at $y^+ = 180$ (dashed lines). $\langle C \rangle_u$ is plotted as a function of the spatial separations in the streamwise (blue) and spanwise directions (red), d_x and d_z, respectively.

To conclude, the mean streamwise velocity profile and the average two-point spatial correlation of the streamwise velocity obtained from the present DNS are shown in Figure B.1(a) and Figure B.1(b), respectively. In particular, the average correlation, $\langle C \rangle_u$, is evaluated at wall-normal positions equal to $y^+ = 3.5$ (i.e., very close to the wall) and $y^+ = 180$ (i.e., in the center of the channel). As illustrated in Figure B.1(b), the extent of the domain in both the homogeneous directions, x and z, is sufficient for the average correlations to decay to zero. Moreover, as shown in Figure B.1(b), the average correlation decays much faster in the spanwise direction than in streamwise direction, namely the (de)correlation scales along z are shorter than along x. This justifies the choice to take a coarser spatial discretization in the streamwise direction than in the spanwise direction.

B.4 Details of the DNS at $Re_\tau = 590$

In this Section we report the details of the direct numerical simulation (DNS) of a turbulent channel flow at $Re_\tau = 590$. The data from this simulation are exploited to build a spatial network from the streamwise velocity field as a comparison with the analysis at $Re_\tau = 180$ (see the complementary results in Chapter 6, Section 6.4).

The numerical approach is the same described in the previous Section B.3, namely the computational techniques and assumptions are the same. In this case, however, a different spatial and temporal discretization is considered. Specifically, the time step used in this simulation equals $1.25 \times 10^{-4} H/u_\tau$, which implies that $\Delta t^+ = \Delta t \cdot u_\tau^2/\nu = 0.07375$ in wall-units. The number of time steps in the simulation is 5017, which corresponds to a time $T u_\tau/H = 0.63$, or $T^+ \approx 370$ in wall-units. In this case, the total temporal window, T^+, is selected to be proportional to the maximum of the integral time-scale (i.e, the integral of the autocorrelation function [4]). The DNS domain sizes, $(L_x, L_y, L_z)_{DNS}$, are equal to $8\pi H$, $2H$ and πH in the streamwise, wall-normal and spanwise directions, (x, y, z), respectively. The number of Fourier modes in the streamwise and spanwise directions equals 512 and 256, respectively, while 257 grid points in the wall-normal direction are exploited. This implies that in physical space the number of simulated grid points equals $(N_x, N_y, N_z)_{DNS} = (768 \times 257 \times 384)$. The DNS grid spacing (in wall-units) in the streamwise and spanwise directions are $\Delta x^+_{DNS} = 19.31$ and $\Delta z^+_{DNS} = 4.83$, respectively. The spacing in the wall-normal direction follows the Eq. (B.3) with $N_y = 257$.

As in the case for $Re_\tau = 180$, the computational domain is reduced in the streamwise and spanwise directions by keeping the DNS discretization, in order to have a manageable network size. In the streamwise direction, $N_x = 192$ out of 768 consecutive grid points are selected, corresponding to a domain length $L_x = L_{x,DNS}/4$. In the spanwise direction, 133 out of 384 consecutive grid points are selected, thus corresponding to a spanwise length $L_z^+ \approx 642$ which is comparable with the $L^+_{z,DNS}$ of the simulation at $Re_\tau = 180$. As for the simulation at $Re_\tau = 180$, the first and last grid points in the wall-normal direction are excluded (due to no-slip condition at the wall). By doing so, the final domain considered for the network analysis is not periodic in the homogeneous directions and the final spatial discretization is $(N_x, N_y, N_z) = (192, 255, 133)$.

B.5 Details of the DNS at $Re_\tau = 950$

In this Section we report the details of the direct numerical simulation (DNS) of a turbulent channel flow at $Re_\tau = 950$, which is exploited to build the temporal network from fluid particle trajectories (see the results in Chapter 7).

The numerical approach is the same described in the previous Section B.3,

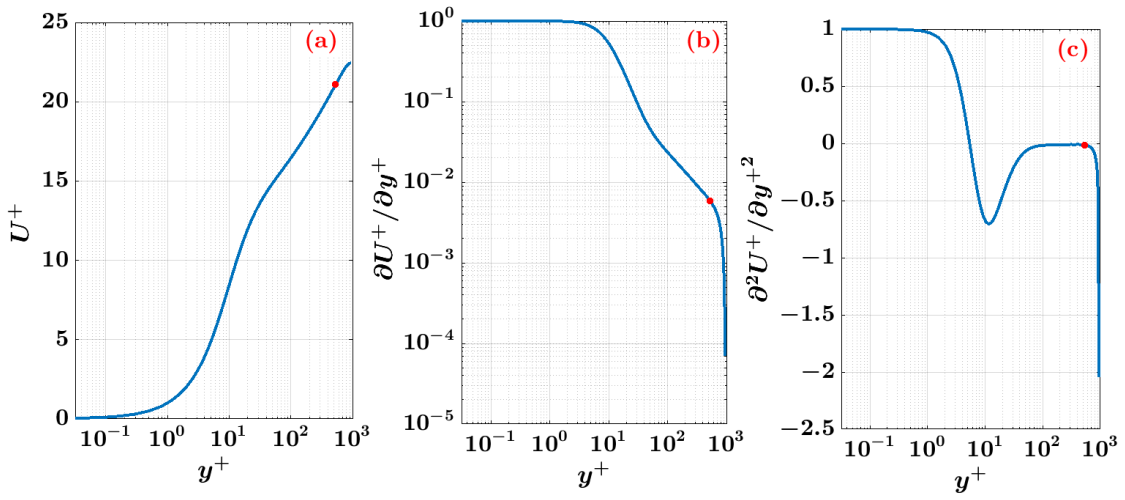


Figure B.2: Wall-normal profiles in wall-units of (a) the mean velocity, U^+, (b) the first derivative of U^+ (i.e., the mean shear), and (c) the second derivative of U^+. The red points highlight the value of each quantity at $y^+ \approx 530$.

namely the computational techniques and assumptions are the same (see Eq. (B.1) and (B.2)). In this case, however, a different spatial and temporal discretization is considered [65]. Here, for $Re_\tau = 950$, 768 Fourier modes are used in the two periodic directions, x and z, while 385 Chebyshev polynomials are employed in the wall-normal direction. The time step (in wall-units) used in the simulation equals $\Delta^+_{t,DNS} = 0.095$, while the grid spacing are $\Delta x^+_{DNS} = 7.8$, $\Delta z^+_{DNS} = 3.9$ and $\Delta y^+_{max} \approx 7.8$ (at the channel center). To characterize the velocity field, Figure B.2 shows the wall-normal profiles (in wall-units) of the mean velocity, U^+, and its first and second derivative. Specifically, the red points highlight the wall-normal coordinate $y^+ \approx 530$, which corresponds to the end of the log-law profile and the beginning of the wake region.

In order to extract the trajectory of a generic particle i, $\boldsymbol{X}_i^+(t^+)$, the equation $\mathrm{d}\boldsymbol{X}_i^+/\mathrm{d}t^+ = \boldsymbol{u}\left(\boldsymbol{X}_i^+(t^+), t^+\right)$, is solved with the same explicit second-order accurate Runge-Kutta method as used in the solution of the NS equations (see Eq. (B.2)). The fluid velocity is interpolated to the particle location, $\boldsymbol{X_i} = \{X_i, Y_i, Z_i\}$, by tri-linear interpolation. The accuracy of the numerical method has been assessed in previous papers (see Ref.s [65, 191]). Figure B.3 shows the average pairwise particle distance in wall-units along the three Cartesian directions, $\overline{d}^+_{x,y,z}$ and in modulus, $\overline{d}^+$, as a function of time. For short times, the largest particle distances are along the transversal directions, y and z, while for long times (especially due to the mean flow advection) the longitudinal component, $\overline{d}^+_x$, dominates. It is worth noting that the average distance in the wall-normal direction, $\overline{d}^+_y$, is always almost constant in time and equal to a third of the channel height (i.e., $2H/3$ or $2Re_\tau/3$

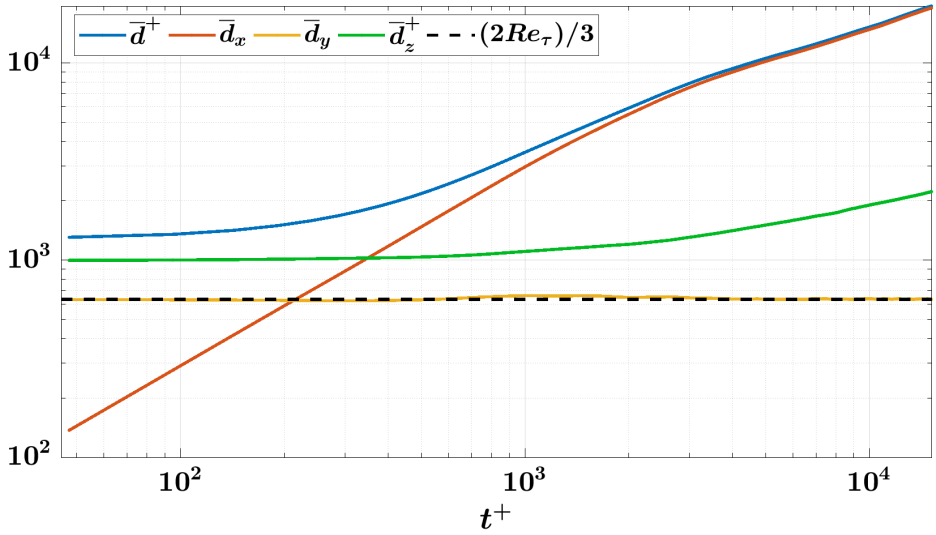


Figure B.3: Average pairwise particle distance in wall-units as a function of time. The dashed black line indicates one third of the channel height.

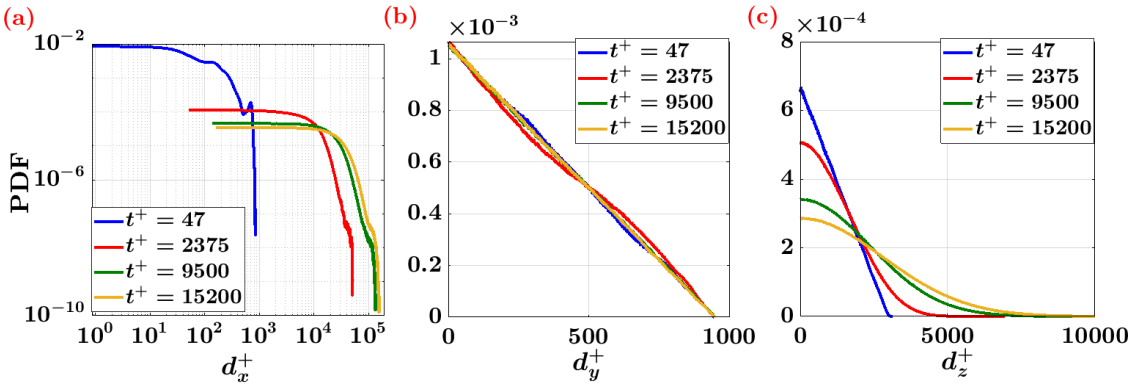


Figure B.4: Probability density function (PDF) at different times, t^+, of the pairwise particle distance in the three Cartesian directions: (a) the streamwise direction, x, (b) the wall-normal direction, y, and (c) the spanwise direction, z. Distances are reported in wall units.

in wall-units). This is due to the fact that, as shown in Figure B.4, while the PDFs of d_x^+ (Figure B.4(a)) and d_z^+ (Figure B.4(c)) significantly change due to particle dispersion, in the wall-normal direction (because of the walls constraint) a linear PDF for d_y^+ (Figure B.4(b)) is always found. This implies that the average value $\overline{d}_y^+$ does not depend on time and equals one third of the maximum distance (equal to $2Re_\tau$ in wall-units), as follows from a triangular distribution.

Appendix C

Experimental measurements on turbulent boundary layer

Some of the contents presented in this Appendix have been previously published. Below are provided the references:

G. Iacobello, M. Marro, L. Ridolfi, P. Salizzoni, and S. Scarsoglio. "Experimental investigation of vertical turbulent transport of a passive scalar in a boundary layer: Statistics and visibility graph analysis". In: *Physical Review Fluids* 4.10 (2019), p. 104501. DOI: 10.1103/PhysRevFluids.4.104501

C.1 Introduction

This Appendix contains details about the experimental setup employed to study the dynamics of a passive scalar plume in a turbulent boundary layer. Specifically, the wind tunnel characteristics and the features of the passive scalar source are reported in Section C.2. The instrument details, the measurement procedure as well as the data pre-processing are described in Section C.3. The results of the network-based analysis of time-series of concentration a vertical turbulent transport are reported in Chapter 5, Section 5.4.

C.2 Description of the wind tunnel setup

A neutrally-stratified atmospheric turbulent boundary layer (TBL) is generated in a recirculating wind tunnel of the Laboratoire de Mécanique des Fluides et d'Acoustique at the École Centrale de Lyon, in France. The setup and the measurement tools are the same as that adopted by *Nironi et al.* [71]. However, measurements are performed in a wind tunnel which smaller than that used by *Nironi et al.* [71], with a working section that is 9 m long, 1 m wide, and 0.7 m

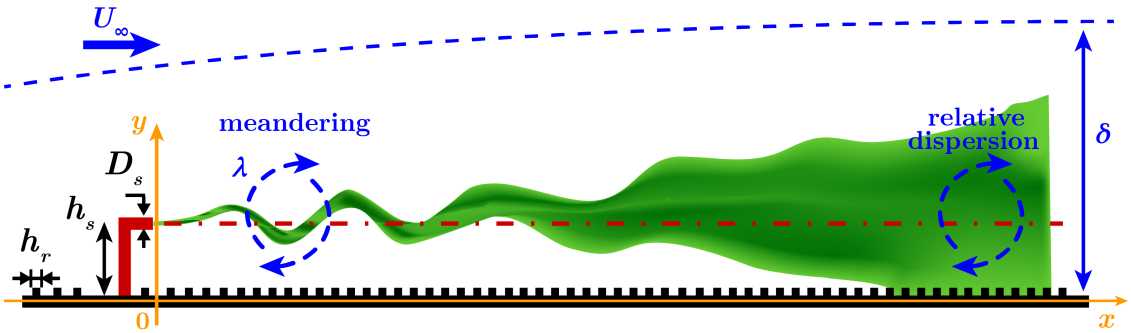


Figure C.1: 2D sketch of the TBL setup in a $(x - y)$ plane [95]. The plume is illustrated in green, while the horizontal dash-dot line refers to the source axis. The symbols are defined in the main text. Two large-scale eddies of (Eulerian) characteristic size λ are also depicted as rotating arrows.

high. A row of Irwin spires [198] is placed at the beginning of the test section, while cubic roughness elements with size $h_r = 0.02$ m are uniformly displaced on the floor. As a result, a TBL of free-stream velocity $U_\infty = 4.94$ m/s and thickness $\delta = 0.314$ m is generated, with δ evaluated as the wall-normal coordinate where the mean velocity $U = 0.95U_\infty$. A sketch of the setup is illustrated in Figure C.1. The Reynolds number of the experiment is evaluated as $Re_\delta = \delta U_\infty/\nu \approx 1.034 \times 10^5$ ($\nu = 1.5 \times 10^{-5}$ m^2/s is the kinematic viscosity of air), which guarantees a well-developed rough turbulent flow [15]. The streamwise, wall-normal and spanwise directions are indicated as (x, y, z), respectively, and the origin of the axes is at the wall in correspondence to the outlet section of the source (see the sketch in Figure C.1).

A mixture of air and a passive scalar is continuously ejected from a metallic L-shaped tube. Due to its density similar to air, Ethane (C_2H_6) is used as a passive tracer. The passive scalar source is located at a streamwise distance from the beginning of the working section approximatively equal to 17.5δ and at a wall-normal height $h_s/\delta \approx 0.24$. Two internal diameter configurations are considered (see Figure C.1): $D_s = 0.003$ m (i.e., $D_s/\delta \approx 9.55 \times 10^{-3}$) and $D_s = 0.006$ m (i.e., $D_s/\delta \approx 1.91 \times 10^{-2}$). The Ethane-air mixture does not substantially introduce or subtract momentum from the flow field at the source. This condition is referred to as *isokinetic* [71, 78], namely the source velocity, U_s, of the mixture equals the local mean velocity, U, at the source height, $U_s \equiv U(y = h_s) \approx 3.37$ m/s. In order to have isokinetic conditions, the total mass flow rate, $M_t = \rho U_s \pi D_s^2/4$, is imposed as $M_t/\rho \approx 86$ l/h for D3 and $M_t/\rho \approx 344$ l/h for D6, where $\rho = \rho_{air} = \rho_{C2H6}$ is the density of the air-Ethane mixture. Furthermore, to consider the recirculation of Ethane-air in the wind tunnel, the background concentration (which increases linearly with time) is subtracted from the recorded time-series.

C.3 Measurement procedures

The streamwise and vertical velocity time-series, u and v, are acquired by means of a X-probe hot-wire anemometer (HWA), while concentration time-series, c, are recorded with a fast Flame Ionization Detector (FID) [199]. The acquisition time is set equal to $T = 180$ s while the number of recorded data is 1.8×10^5. Each time-series is normalized by a reference value, which is U_∞ for the velocity components (measured in m/s) and $\Delta c = M_e/(\rho U_\infty \delta^2)$ for the passive scalar concentration (measured in ppm), where M_e is the mass flow rate of Ethane. To take into account the presence of random instrumental noise on c – that produces negative concentration values – the concentration data are preprocessed as

$$c(x, y, z; t_i) = 0, \quad \text{if} \quad c(x, y, z; t_i) < \theta_c, \tag{C.1}$$

where $\theta_c = |\min_{x,y,z} [\min_{t_i} [c(x, y, z; t_i)]]|$ is the absolute value of the minimum amplitude of all concentration series for a given D_s. In other words, we set equal to zero all concentration values that are smaller (in modulus) than the maximum amplitude of negative values in the series. This pre-processing operation is reasonably valid as the values of θ_c are two orders of magnitude lower than the average concentration values, and three orders of magnitude lower than the maximum c values. Furthermore, the Reynolds decomposition is performed for velocity and concentration time-series as $u' = u - U$, $v' = v - \overline{v}$ and $c' = c - \overline{c}$, where $\overline{v}$ and $\overline{c}$ are the fixed-point time-averages of v and c, respectively. The details of the velocity and concentration measurements are reported as follows.

Velocity measurements

Velocity time-series are acquired by means of a $\pm 45°$ X-probe hot-wire anemometer (HWA) working at a constant temperature, which allows for the simultaneous measurements of two velocity components. The probe is calibrated by exploiting a Pitot tube that measures a reference velocity (calibration in yaw is not performed). In particular, the calibration velocities are decomposed into the longitudinal and transversal velocity components by adopting a yaw correction with constant coefficients $K_1^2 = K_2^2 = 0.0225$ [200]. The experimental error at a fixed reference location is approximately $\pm 2\%$ for the mean and the standard deviation.

The main features of the velocity field are shown in Figure C.2. Specifically, the vertical profile of the mean velocity defect is displayed in Figure C.2(a), while standard deviations of the velocity components, σ_u and σ_v, and the Reynolds stress, $\overline{u'v'}$, are shown in Figure C.2(b) and Figure C.2(c), respectively. The vertical profiles – measured at different streamwise locations in the range $x/\delta \in [0.65, 3.90]$ – are in good agreement with values reported in literature [71], and collapse rather well both for the mean flow (Figure C.2(a)) and the velocity fluctuations (Figure C.2(b),(c)). In particular, from the Reynolds stress profile it is possible to estimate the friction

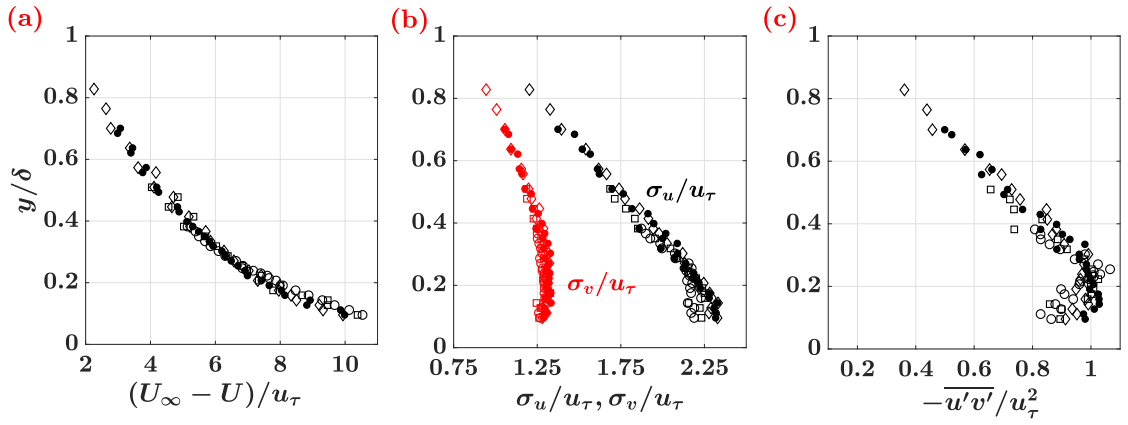


Figure C.2: Vertical profiles of velocity statistics normalized with the friction velocity, u_τ. (a) Mean velocity defect, $(U_\infty - U)$. (b) Standard deviations of the longitudinal, σ_u, and vertical, σ_v, velocity component, depicted in black and red, respectively. (c) Reynolds stress, $-\overline{u'v'}$. Symbols: $x/\delta = 0.65, \circ$; $x/\delta = 1.30, \square$; $x/\delta = 2.60, \bullet$; $x/\delta = 3.90, \lozenge$.

velocity as $u_\tau = \left(-\overline{u'v'}\right)^{0.5} = 0.209$ m/s, by averaging the profiles of $\overline{u'v'}$ in the region close to the wall [71].

Concentration measurements

A fast Flame Ionization Detector (FID) is used to perform concentration measurements. The FID system uses a sampling tube that is 0.3 m long, permitting a frequency response of the instrument to about 400 Hz. The calibration is carried out twice a day by setting ethane-air concentrations equal to 0, 500, 1000 and 5000 ppm. A linear relation holds between ethane concentration and tension response, with slope variations (i.e., the sensitivity variations of the instrument) of about $\pm 3\%$, depending on the ambient conditions. The error in the first four moments of the concentration due to all the uncertainties in the experimental chain, is estimated to be up to 4.5%. By exploiting measurements performed on different days of distant weeks, the first two moments of the concentration are affected by an error of 2% in the far field and 3% in the near field (this increase is due to uncertainties in the source flow control system in the near field). For the third and fourth moments of the concentration, the error rises up to 4.5% both in the near- and far-field. Furthermore, in order to evaluate one-point turbulent fluxes, simultaneous measurements of velocity and concentration are necessary at the same spatial location, implying that HWA and FID systems have to be synchronized and sufficiently close in space. The optimal distance of the FID with respect to the HWA is found to be 5 mm in the spanwise direction, in order to avoid local perturbations induced

by the measuring system in the flow field. The coupling HWA-FID does not require signal re-sampling or filtering, because both HWA and FID have a constant sampling frequency (equal to 1000 Hz), so that their responses are continuous and regular.

Finally, the vertical position, h_s^*, of the actual axis of the plume is estimated for both D3 and D6 as the y coordinate of maximum $\overline{c}(y)$ value. In fact, the maximum value of the mean concentration, h_s^*, along y/δ is not exactly at $y = h_s$, but it decreases downstream from the source mainly due to two factors: the effect of the mean shear, $\partial U/\partial y$, and the source wake. In order to extract a reliable value of the wall-normal coordinate of maximum $\overline{c}$, we fitted the vertical profiles of mean concentration (see Figure 5.13(a)) by adopting a Gaussian distribution with total reflection on the ground [181], defined as

$$\overline{c}_{fit}\left(x,y\right) = \frac{M_e/\rho}{2\pi\Lambda_y\Lambda_z\overline{U}_{adv}}\left[\exp\left(-\frac{\left(y+h_s^*\right)^2}{2\Lambda_y^2}\right)+\exp\left(-\frac{\left(y-h_s^*\right)^2}{2\Lambda_y^2}\right)\right] \tag{C.2}$$

where $\overline{U}_{adv}$ is the mean streamwise velocity at the plume center of mass, h_s^* is the vertical coordinate of the plume axis, while Λ_y and Λ_z are the wall-normal and transversal (average) spread of the plume, respectively. The fitting function reported in Eq. (C.2) is the most suited distribution to reproduce the vertical mean concentration profiles [181]. In particular, here Λ_y, Λ_z and h_s^* are adopted as free parameters of the fitting procedure. Although the value of h_s^* can also be set as constant (as an approximation) and equal to h_s, in this work we explicitly used h_s^* as a free parameter in the Eq. (C.2) to highlight the effect of the mean shear on the plume. By doing so, h_s^* is not fixed but depends on the source size by moving downstream, with $h_s^* \leq h_s$. In Figure 5.13, the values of h_s^* are shown as horizontal dashed and dot-dashed lines for D3 and D6, respectively. It is also worth noting that the values of the vertical coordinate of the plume axis, h_s^*, are better identified by the parameters of the fitting procedure than by locating the maximum (experimental) $\overline{c}$ value. This issue is crucial in the far field (i.e., far downstream from the source), where $h_s^* > 0$ whereas the maximum (experimental) $\overline{c}$ value is located very close to the floor (i.e., $(y/\delta)_{\overline{c}_{max}} \to 0$, due to the plume dispersion and the effect of the ground reflection of the plume).

Bibliography

[1] H. Tennekes and J. L. Lumley. *A first course in turbulence*. MIT press, (1972). URL: mitpress.mit.edu/books/first-course-turbulence.

[2] U. Frisch. *Turbulence: the legacy of AN Kolmogorov*. Cambridge University press, (1995). DOI: 10.1017/CBO9781139170666.

[3] J. Mathieu and J. Scott. *An introduction to turbulent flow*. Cambridge University Press, (2000). DOI: 10.1017/CBO9781316529850.

[4] S. B. Pope. *Turbulent flows*. Cambridge University press, (2001). DOI: 10.1017/CBO9780511840531.

[5] P. Davidson. *Turbulence: an introduction for scientists and engineers*. Oxford University Press, (2015). DOI: 10.1093/acprof:oso/9780198722588.001.0001.

[6] J. C. Wyngaard. *Turbulence in the Atmosphere*. Cambridge University Press, (2010). DOI: 10.1017/CBO9780511840524.

[7] B. Sawford. "Turbulent relative dispersion". In: *Annual Review of Fluid Mechanics* 33.1 (2001), pp. 289–317. DOI: 10.1146/annurev.fluid.33.1.289.

[8] P. E. Dimotakis. "Turbulent mixing". In: *Annual Review of Fluid Mechanics* 37 (2005), pp. 329–356. DOI: 10.1146/annurev.fluid.36.050802.122015.

[9] G. Boffetta and R. E. Ecke. "Two-dimensional turbulence". In: *Annual Review of Fluid Mechanics* 44 (2012), pp. 427–451. DOI: 10.1146/annurev-fluid-120710-101240.

[10] T. Günther, A. Kuhn, H.C. Hege, M. Gross, and H. Theisel. "Progressive Monte Carlo rendering of atmospheric flow features across scales". In: *Physical Review Fluids* 2.9 (2017), p. 090502. DOI: 10.1103/PhysRevFluids.2.090502.

[11] I. Procaccia and K. R. Sreenivasan. "The state of the art in hydrodynamic turbulence: Past successes and future challenges". In: *Physica D: Nonlinear Phenomena* 237.14-17 (2008), pp. 2167–2183. DOI: 10.1016/j.physd.2008.01.025.

[12] J. Jiménez. "Near-wall turbulence". In: *Physics of Fluids* 25.10 (2013), p. 101302. DOI: 10.1063/1.4824988.

[13] A. J. Smits and I. Marusic. "Wall-bounded turbulence". In: *Physics Today* 66.9 (2013), pp. 25–30. DOI: 10.1063/PT.3.2114.

[14] I. Marusic and R. J. Adrian. "The Eddies and Scales of Wall Turbulence". In: *Ten Chapters in Turbulence*. Ed. by P. A. Davidson, Y. Kaneda, and K. R. Sreenivasan. Cambridge University Press, (2012.

[15] J. Jiménez. "Turbulent flows over rough walls". In: *Annual Review of Fluid Mechanics* 36 (2004), pp. 173–196. DOI: 10.1146/annurev.fluid.36.050802.122103.

[16] A. J. Smits, B. J McKeon, and I. Marusic. "High–Reynolds number wall turbulence". In: *Annual Review of Fluid Mechanics* 43 (2011), pp. 353–375. DOI: 10.1146/annurev-fluid-122109-160753.

[17] O. Reynolds. "On the dynamical theory of incompressible viscous fluids and the determination of the criterion". In: *Proceedings of the Royal Society of London* 56.336-339 (1894), pp. 40–45. DOI: 10.1098/rspl.1894.0075.

[18] G. I. Taylor. "Diffusion by continuous movements". In: *Proceedings of the London Mathematical Society* 2.1 (1922), pp. 196–212. DOI: 10.1112/plms/s2-20.1.196.

[19] G. I. Taylor. "Statistical Theory of Turbulence". In: *Proceedings of the Royal Society of London A* 151.873 (1935), pp. 421–444. DOI: 10.1098/rspa.1935.0158.

[20] A. N. Kolmogorov. "The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers". In: *Proceedings of the Royal Society of London A* 434.1890 (1991), pp. 9–13. DOI: 10.1098/rspa.1991.0075.

[21] A. N. Kolmogorov. "Dissipation of energy in the locally isotropic turbulence". In: *Proceedings of the Royal Society of London A* 434.1890 (1991), pp. 15–17. DOI: 10.1098/rspa.1991.0076.

[22] T. Kajishima and K. Taira. *Computational fluid dynamics*. Springer, (2017). DOI: 10.1007/978-3-319-45304-0.

[23] P. Moin and K. Mahesh. "Direct numerical simulation: a tool in turbulence research". In: *Annual Review of Fluid Mechanics* 30.1 (1998), pp. 539–578. DOI: 10.1146/annurev.fluid.30.1.539.

[24] S. J. Kline, W. C. Reynolds, F.A. Schraub, and P.W. Runstadler. "The structure of turbulent boundary layers". In: *Journal of Fluid Mechanics* 30.4 (1967), pp. 741–773. DOI: 10.1017/S0022112067001740.

[25] E. R. Corino and R. S. Brodkey. "A visual investigation of the wall region in turbulent flow". In: *Journal of Fluid Mechanics* 37.1 (1969), pp. 1–30. DOI: 10.1017/S0022112069000395.

[26] S. A. Orszag and G.S. Patterson Jr. "Numerical simulation of three-dimensional homogeneous isotropic turbulence". In: *Physical Review Letters* 28.2 (1972), p. 76. DOI: 10.1103/PhysRevLett.28.76.

[27] J. Kim, P. Moin, and R. Moser. "Turbulence statistics in fully developed channel flow at low Reynolds number". In: *Journal of Fluid Mechanics* 177 (1987), pp. 133–166. DOI: 10.1017/S0022112087000892.

[28] J N. Kutz. "Deep learning in fluid dynamics". In: *Journal of Fluid Mechanics* 814 (2017), pp. 1–4. DOI: 10.1017/jfm.2016.803.

[29] K. Fukami, K. Fukagata, and K. Taira. "Super-resolution reconstruction of turbulent flows with machine learning". In: *Journal of Fluid Mechanics* 870 (2019), pp. 106–120. DOI: 10.1017/jfm.2019.238.

[30] S. L. Brunton, B. R. Noack, and P. Koumoutsakos. "Machine learning for fluid mechanics". In: *Annual Review of Fluid Mechanics* 52 (2019). DOI: 10.1146/annurev-fluid-010719-060214.

[31] A. Pollard. "Whither Turbulence and Big Data for the Twenty-First Century". In: *Pollard A., Castillo L., Danaila L., Glauser M. (eds) Whither Turbulence and Big Data in the 21st Century?* Springer, (2017), pp. 531–547. DOI: 10.1007/978-3-319-41217-7.

[32] J. Jiménez. "Coherent structures in wall-bounded turbulence". In: *Journal of Fluid Mechanics* 842 (2018). DOI: 10.1017/jfm.2018.144.

[33] B.J. McKeon and K.R Sreenivasan. "Introduction: scaling and structure in high Reynolds number wall-bounded flows". In: *Philosophical Transactions of the Royal Society A* 365.1852 (2007), pp. 635–646. DOI: 10.1098/rsta.2006.1952.

[34] P. Krogstadt and R.A. Antonia. "Surface roughness effects in turbulent boundary layers". In: *Experiments in Fluids* 27.5 (1999), pp. 450–460. DOI: 10.1007/s003480050370.

[35] S. Heinz. "On mean flow universality of turbulent wall flows. I. High Reynolds number flow analysis". In: *Journal of Turbulence* 19.11-12 (2018), pp. 929–958. DOI: 10.1080/14685248.2019.1566736.

[36] I. Marusic, B.J. McKeon, P. A. Monkewitz, H.M. Nagib, A.J Smits, and K.R. Sreenivasan. "Wall-bounded turbulent flows at high Reynolds numbers: recent advances and key issues". In: *Physics of Fluids* 22.6 (2010), p. 065103. DOI: 10.1063/1.3453711.

[37] S. Heinz. "On mean flow universality of turbulent wall flows. II. Asymptotic flow analysis". In: *Journal of Turbulence* (2019), pp. 1–20. DOI: 10.1080/14685248.2019.1593425.

[38] P. Luchini. "Universality of the turbulent velocity profile". In: *Physical Review Letters* 118.22 (2017), p. 224501. DOI: 10.1103/PhysRevLett.118.224501.

[39] R. J. Adrian. "Hairpin vortex organization in wall turbulence". In: *Physics of Fluids* 19.4 (2007), p. 041301. DOI: 10.1063/1.2717527.

[40] E. Perlman, R. Burns, Y. Li, and C. Meneveau. "Data exploration of turbulence simulations using a database cluster". In: *Proc. of the 2007 ACM/IEEE Conf. on Supercomputing*. ACM. (2007), p. 23. DOI: 10.1145/1362622.1362654.

[41] J. M Wallace. "Quadrant analysis in turbulence research: history and evolution". In: *Annual Review of Fluid Mechanics* 48 (2016), pp. 131–158. DOI: 10.1146/annurev-fluid-122414-034550.

[42] A. Lozano-Durán, O. Flores, and J. Jiménez. "The three-dimensional structure of momentum transfer in turbulent channels". In: *Journal of Fluid Mechanics* 694 (2012), pp. 100–130. DOI: 10.1017/jfm.2011.524.

[43] P. Chakraborty, S. Balachandar, and R. J Adrian. "On the relationships between local vortex identification schemes". In: *Journal of Fluid Mechanics* 535 (2005), pp. 189–214. DOI: 10.1017/S0022112005004726.

[44] J.A. Sillero, J. Jiménez, and R.D. Moser. "Two-point statistics for turbulent boundary layers and channels at Reynolds numbers up to $\delta^+ \approx$". In: *Physics of Fluids* 26.10 (2014), p. 105109. DOI: 10.1063/1.4899259.

[45] N. Hutchins and I. Marusic. "Large-scale influences in near-wall turbulence". In: *Philosophical Transactions of the Royal Society A* 365.1852 (2007), pp. 647–664. DOI: 10.1098/rsta.2006.1942.

[46] J. Hwang, J. Lee, H. J. Sung, and T. A. Zaki. "Inner–outer interactions of large-scale structures in turbulent channel flow". In: *Journal of Fluid Mechanics* 790 (2016), pp. 128–157. DOI: 10.1017/jfm.2016.3.

[47] M.F. Howland and X.I.A. Yang. "Dependence of small-scale energetics on large scales in turbulent flows". In: *Journal of Fluid Mechanics* 852 (2018), pp. 641–662. DOI: 10.1017/jfm.2018.554.

[48] G. Berkooz, P. Holmes, and J. L. Lumley. "The proper orthogonal decomposition in the analysis of turbulent flows". In: *Annual Review of Fluid Mechanics* 25.1 (1993), pp. 539–575. DOI: 10.1017/S0022112005004726.

[49] R. J. Adrian and P. Moin. "Stochastic estimation of organized turbulent structure: homogeneous shear flow". In: *Journal of Fluid Mechanics* 190 (1988), pp. 531–559. DOI: 10.1017/S0022112088001442.

[50] J. Chen, F. Hussain, J. Pei, and Z.S. She. "Velocity–vorticity correlation structure in turbulent channel flow". In: *Journal of Fluid Mechanics* 742 (2014), pp. 291–307. DOI: 10.1017/jfm.2014.3.

[51] Z. Warhaft. "Passive scalars in turbulent flows". In: *Annual Review of Fluid Mechanics* 32.1 (2000), pp. 203–240. DOI: 10.1146/annurev.fluid.32.1.203.

[52] L. F. Richardson. "Atmospheric diffusion shown on a distance-neighbour graph". In: *Proceedings of the Royal Society of London A* 110.756 (1926), pp. 709–737. DOI: 10.1098/rspa.1926.0043.

[53] K. R. Sreenivasan and J. Schumacher. "Lagrangian views on turbulent mixing of passive scalars". In: *Philosophical Transactions of the Royal Society A* 368.1916 (2010), pp. 1561–1577. DOI: 10.1098/rsta.2009.0140.

[54] G. Falkovich, K. Gawedzki, and M. Vergassola. "Particles and fields in fluid turbulence". In: *Reviews of Modern Physics* 73.4 (2001), p. 913. DOI: 10.1103/RevModPhys.73.913.

[55] F. Toschi and E. Bodenschatz. "Lagrangian properties of particles in turbulence". In: *Annual Review of Fluid Mechanics* 41 (2009), pp. 375–404. DOI: 10.1146/annurev.fluid.010908.165210.

[56] J. Salazar and L. R Collins. "Two-particle dispersion in isotropic turbulent flows". In: *Annual Review of Fluid Mechanics* 41 (2009), pp. 405–432. DOI: 10.1146/annurev.fluid.40.111406.102224.

[57] A. Pumir, B.I. Shraiman, and M. Chertkov. "Geometry of Lagrangian dispersion in turbulence". In: *Physical Review Letters* 85.25 (2000), p. 5324. DOI: doi.org/10.1103/PhysRevLett.85.5324.

[58] L. Biferale, G. Boffetta, A. Celani, B.J. Devenish, A. Lanotte, and F. Toschi. "Multiparticle dispersion in fully developed turbulence". In: *Physics of Fluids* 17.11 (2005), p. 111701. DOI: doi.org/10.1063/1.2130751.

[59] B. Lüthi, S. Ott, J. Berg, and J. Mann. "Lagrangian multi-particle statistics". In: *Journal of Turbulence* 8 (2007), N45. DOI: doi.org/10.1080/14685240701522927.

[60] J. Schumacher. "Lagrangian studies in convective turbulence". In: *Physical Review E* 79.5 (2009), p. 056301. DOI: doi.org/10.1103/PhysRevE.79.056301.

[61] S. Bianchi, L. Biferale, A. Celani, and M. Cencini. "On the evolution of particle-puffs in turbulence". In: *European Journal of Mechanics-B/Fluids* 55 (2016), pp. 324–329. DOI: doi.org/10.1016/j.euromechflu.2015.06.009.

[62] J. I. Polanco, I. Vinkovic, N. Stelzenmuller, N. Mordant, and M. Bourgoin. "Relative dispersion of particle pairs in turbulent channel flow". In: *International Journal of Heat and Fluid Flow* 71 (2018), pp. 231–245. DOI: 10.1016/j.ijheatfluidflow.2018.04.007.

[63] J. I. Polanco. "Lagrangian properties of turbulent channel flow: a numerical study". PhD thesis. Université Claude Bernard Lyon 1; Université de Lyon, 2019. URL: https://hal.archives-ouvertes.fr/tel-02084215.

[64] E. Pitton, C. Marchioli, V. Lavezzo, A. Soldati, and F. Toschi. "Anisotropy in pair dispersion of inertial particles in turbulent channel flow". In: *Physics of Fluids* 24.7 (2012), p. 073305. DOI: 10.1063/1.4737655.

[65] J.G.M. Kuerten and J.J.H. Brouwers. "Lagrangian statistics of turbulent channel flow at $Re_\tau = 950$ calculated with direct numerical simulation and Langevin models". In: *Physics of Fluids* 25.10 (2013), p. 105108. DOI: 10.1063/1.4824795.

[66] N. Stelzenmuller, J. I. Polanco, L. Vignal, I. Vinkovic, and N. Mordant. "Lagrangian acceleration statistics in a turbulent channel flow". In: *Physical Review Fluids* 2.5 (2017), p. 054602. DOI: 10.1103/PhysRevFluids.2.054602.

[67] P.C. Chatwin and P. J. Sullivan. "A simple and unifying physical interpretation of scalar fluctuation measurements from many turbulent shear flows". In: *Journal of Fluid Mechanics* 212 (1990), pp. 533–556. DOI: 10.1017/S0022112090002087.

[68] N. Mole and E.D. Clarke. "Relationships between higher moments of concentration and of dose in turbulent dispersion". In: *Boundary-Layer Meteorology* 73.1-2 (1995), pp. 35–52. DOI: 10.1007/BF00708929.

[69] E. Yee and R. Chan. "A simple model for the probability density function of concentration fluctuations in atmospheric plumes". In: *Atmospheric Environment* 31.7 (1997), pp. 991–1002. DOI: 10.1016/S1352-2310(96)00255-5.

[70] E. Villermaux and J. Duplat. "Mixing as an aggregation process". In: *Physical Review Letters* 91.18 (2003), p. 184501. DOI: 10.1016/S1631-0721(03)00110-4.

[71] C. Nironi, P. Salizzoni, M. Marro, P. Mejean, N. Grosjean, and L. Soulhac. "Dispersion of a passive scalar fluctuating plume in a turbulent boundary layer. Part I: Velocity and concentration measurements". In: *Boundary-Layer meteorology* 156.3 (2015), pp. 415–446. DOI: 10.1007/s10546-015-0040-x.

[72] E. Yee, R. Chan, P.R. Kosteniuk, G.M. Chandler, C.A. Biltoft, and J.F. Bowers. "Incorporation of internal fluctuations in a meandering plume model of concentration fluctuations". In: *Boundary-Layer Meteorology* 67.1-2 (1994), pp. 11–39. DOI: 10.1007/BF00705506.

[73] P. Franzese. "Lagrangian stochastic modeling of a fluctuating plume in the convective boundary layer". In: *Atmospheric Environment* 37.12 (2003), pp. 1691–1701. DOI: 10.1016/S1352-2310(03)00003-7.

[74] M. Marro, C. Nironi, P. Salizzoni, and L. Soulhac. "Dispersion of a passive scalar fluctuating plume in a turbulent boundary layer. part II: Analytical modelling". In: *Boundary-Layer meteorology* 156.3 (2015), pp. 447–469. DOI: 10.1007/s10546-015-0041-9.

[75] A. Bisignano, L. Mortarini, and E. Ferrero. "Evaluation of high-order concentration statistics in a dispersing plume". In: *Physica A: Statistical Mechanics and its Applications* 474 (2017), pp. 115–126. DOI: 10.1016/j.physa.2017.01.068.

[76] F. Gifford Jr. "Statistical properties of a fluctuating plume dispersion model". In: *Advances in geophysics*. Vol. 6. Elsevier, (1959), pp. 117–137. DOI: 10.1016/S0065-2687(08)60099-0.

[77] J.E. Fackrell and A.G. Robins. "Concentration fluctuations and fluxes in plumes from point sources in a turbulent boundary layer". In: *Journal of Fluid Mechanics* 117 (1982), pp. 1–26. DOI: 10.1017/S0022112082001499.

[78] K.M. Talluru, J. Philip, and K.A. Chauhan. "Local transport of passive scalar released from a point source in a turbulent boundary layer". In: *Journal of Fluid Mechanics* 846 (2018), pp. 292–317. DOI: 10.1017/jfm.2018.280.

[79] K.M. Talluru, J. Philip, and K.A. Chauhan. "Self-similar spectra of point-source scalar plumes in a turbulent boundary layer". In: *Journal of Fluid Mechanics* 870 (2019), pp. 698–717. DOI: 10.1017/jfm.2019.253.

[80] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.U. Hwang. "Complex networks: Structure and dynamics". In: *Physics Reports* 424.4-5 (2006), pp. 175–308. DOI: 10.1016/j.physrep.2005.10.009.

[81] D. J. Watts and S. H. Strogatz. "Collective dynamics of 'small-world'networks". In: *Nature* 393.6684 (1998), p. 440. DOI: 10.1038/30918.

[82] A.L. Barabási and R. Albert. "Emergence of scaling in random networks". In: *Science* 286.5439 (1999), pp. 509–512. DOI: 10.1126/science.286.5439.509.

[83] K. A. Zweig. *Network analysis literacy: a practical approach to the analysis of networks*. Springer Science & Business Media, (2016). DOI: 10.1007/978-3-7091-0741-6.

[84] M. Newman. *Networks*. 2nd ed. Oxford University Press, (2018). DOI: 10.1093/oso/9780198805090.001.0001.

[85] L. da F. Costa, F. A. Rodrigues, G. Travieso, and P. R. Villas Boas. "Characterization of complex networks: A survey of measurements". In: *Advances in Physics* 56.1 (2007), pp. 167–242. DOI: 10.1080/00018730601170527.

[86] K. Goh, M.E. Cusick, D. Valle, B. Childs, M. Vidal, and A.L. Barabási. "The human disease network". In: *Proceedings of the National Academy of Sciences* 104.21 (2007), pp. 8685–8690. DOI: 10.1073/pnas.0701361104.

[87] L. da F. Costa, O. N. Oliveira Jr, G. Travieso, F. A. Rodrigues, P. R. Villas Boas, L. Antiqueira, M. P. Viana, and L. E. Correa Rocha. "Analyzing and modeling real-world phenomena with complex networks: a survey of applications". In: *Advances in Physics* 60.3 (2011), pp. 329–412. DOI: 10.1080/00018732.2011.572452.

[88] F. Menczer. "Misinformation on social media: Can technology save us". In: *The Conversation* 27 (2016). URL: http://theconversation.com/misinformation-on-social-media-can-technology-save-us-69264.

[89] F. Silva and L. da F. Costa. "Visualizing Complex Networks (CDT-5)". In: *Costa's Didactic Texts - CDTs*. Nov. 2018. DOI: 10.13140/RG.2.2.21310.74567/1.

[90] G. Iacobello, S. Scarsoglio, J.G.M. Kuerten, and L. Ridolfi. "Spatial characterization of turbulent channel flow via complex networks". In: *Physical Review E* 98.1 (2018), p. 013107. DOI: 10.1103/PhysRevE.98.013107.

[91] J. Heitzig, J. F. Donges, Y. Zou, N. Marwan, and J. Kurths. "Node-weighted measures for complex networks with spatially embedded, sampled, or differently sized nodes". In: *The European Physical Journal B* 85.1 (2012), pp. 1–22. DOI: 10.1140/epjb/e2011-20678-7.

[92] P. Holme and J. Saramäki. "Temporal networks". In: *Physics Reports* 519.3 (2012), pp. 97–125. DOI: 10.1016/j.physrep.2012.03.001.

[93] A. Barrat, M. Barthélemy, and A. Vespignani. "Weighted evolving networks: coupling topology and weight dynamics". In: *Physical Review Letters* 92.22 (2004), p. 228701. DOI: 10.1103/PhysRevLett.92.228701.

[94] G. Iacobello, S. Scarsoglio, and L. Ridolfi. "Visibility graph analysis of wall turbulence time-series". In: *Physics Letters A* 382.1 (2018), pp. 1–11. DOI: 10.1016/j.physleta.2017.10.027.

[95] G. Iacobello, M. Marro, L. Ridolfi, P. Salizzoni, and S. Scarsoglio. "Experimental investigation of vertical turbulent transport of a passive scalar in a boundary layer: Statistics and visibility graph analysis". In: *Physical Review Fluids* 4.10 (2019), p. 104501. DOI: 10.1103/PhysRevFluids.4.104501.

[96] R.V. Donner, E. Hernández-García, and E. Ser-Giacomi. *Introduction to Focus Issue: Complex network perspectives on flow systems*. 2017. DOI: 10.1063/1.4979129.

[97] L. Lacasa, B. Luque, F. Ballesteros, J. Luque, and J. C. Nuno. "From time series to complex networks: The visibility graph". In: *Proceedings of the National Academy of Sciences* 105.13 (2008), pp. 4972–4975. DOI: 10.1073/pnas.0709247105.

[98] Z. Gao and N. Jin. "Complex network from time series based on phase space reconstruction". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 19.3 (2009), p. 033137. DOI: 10.1063/1.3227736.

[99] Z. Gao and N.D. Jin. "Flow-pattern identification and nonlinear dynamics of gas-liquid two-phase flow in complex networks". In: *Physical Review E* 79.6 (2009), p. 066303. DOI: 10.1103/PhysRevE.79.066303.

[100] Z. Gao, X.W. Zhang, N.D. Jin, R. V. Donner, N. Marwan, and J. Kurths. "Recurrence networks from multivariate signals for uncovering dynamic transitions of horizontal oil-water stratified flows". In: *EPL (Europhysics Letters)* 103.5 (2013), p. 50004. DOI: 10.1209/0295-5075/103/50004.

[101] Z. Gao, Y.X.n Yang, L.S. Zhai, M.S. Ding, and N.D. Jin. "Characterizing slug to churn flow transition by using multivariate pseudo Wigner distribution and multivariate multiscale entropy". In: *Chemical Engineering Journal* 291 (2016), pp. 74–81. DOI: 10.1016/j.cej.2016.01.039.

[102] W. Ren and N. Jin. "Vector visibility graph from multivariate time series: a new method for characterizing nonlinear dynamic behavior in two-phase flow". In: *Nonlinear Dynamics* (2019), pp. 1–10. DOI: 10.1007/s11071-019-05147-7.

[103] A.H. Shirazi, G. R. Jafari, J. Davoudi, J. Peinke, M. R. R. Tabar, and M. Sahimi. "Mapping stochastic processes onto complex networks". In: *Journal of Statistical Mechanics: Theory and Experiment* 2009.07 (2009), P07046. DOI: 10.1088/1742-5468/2009/07/P07046.

[104] A.K. Charakopoulos, T.E. Karakasidis, P.N. Papanicolaou, and A. Liakopoulos. "The application of complex network time series analysis in turbulent heated jets". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 24.2 (2014), p. 024408. DOI: 10.1063/1.4875040.

[105] P. Manshour, M. R. R. Tabar, and J. Peinke. "Fully developed turbulence in the view of horizontal visibility graphs". In: *Journal of Statistical Mechanics: Theory and Experiment* 2015.8 (2015), P08031. DOI: 10.1088/1742-5468/2015/08/P08031.

[106] M. Murugesan, Y. Zhu, and L. K.B. Li. "Complex network analysis of forced synchronization in a hydrodynamically self-excited jet". In: *International Journal of Heat and Fluid Flow* 76 (2019), pp. 14–25. DOI: 10.1016/j.ijheatfluidflow.2018.12.004.

[107] G. M. Ramirez Ávila, A. Gapelyuk, N. Marwan, T. Walther, H. Stepan, J. Kurths, and N. Wessel. "Classification of cardiovascular time series based on different coupling structures using recurrence networks analysis". In: *Philosophical Transactions of the Royal Society A* 371.1997 (2013), p. 20110623. DOI: 10.1098/rsta.2011.0623.

[108] S. Scarsoglio, F. Cazzato, and L. Ridolfi. "From time-series to complex networks: Application to the cerebrovascular flow patterns in atrial fibrillation". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 27.9 (2017), p. 093107. DOI: 10.1063/1.5003791.

[109] A.C. Braga, L.G.A. Alves, L.S. Costa, A.A. Ribeiro, M.M.A. de Jesus, A.A. Tateishi, and H.V. Ribeiro. "Characterization of river flow fluctuations via horizontal visibility graphs". In: *Physica A: Statistical Mechanics and its Applications* 444 (2016), pp. 1003–1011. DOI: 10.1016/j.physa.2015.10.102.

[110] X. Han, B. Sivakumar, F. M. Woldmeskel, and M. G. de Aguilar. "Temporal dynamics of streamflow: application of complex networks". In: *Geoscience Letters* 5.1 (2018), p. 10. DOI: 10.1186/s40562-018-0109-8.

[111] A.K. Charakopoulos, G.A. Katsouli, and T.E. Karakasidis. "Dynamics and causalities of atmospheric and oceanic data identified by complex networks and Granger causality analysis". In: *Physica A: Statistical Mechanics and its Applications* 495 (2018), pp. 436–453. DOI: 10.1016/j.physa.2017.12.027.

[112] A. Charakopoulos, T. Karakasidis, and I. Sarris. "Pattern identification for wind power forecasting via complex network and recurrence plot time series analysis". In: *Energy Policy* 133 (2019), p. 110934. DOI: 10.1016/j.enpol.2019.110934.

[113] M. Murugesan and R.I. Sujith. "Combustion noise is scale-free: transition from scale-free to order at the onset of thermoacoustic instability". In: *Journal of Fluid Mechanics* 772 (2015), pp. 225–245. DOI: 10.1017/jfm.2015.215.

[114] V. Godavarthi, V.R. Unni, E.A. Gopalakrishnan, and R.I. Sujith. "Recurrence networks to study dynamical transitions in a turbulent combustor". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 27.6 (2017), p. 063113. DOI: 10.1063/1.4985275.

[115] J. Singh, R. Belur Vishwanath, S. Chaudhuri, and R.I. Sujith. "Network structure of turbulent premixed flames". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 27.4 (2017), p. 043107. DOI: 10.1063/1.4980135.

[116] H. Zan, W. Zhou, X. Xiao, L. Lin, J. Zhang, and H. Li. "Recurrence network analysis for uncovering dynamic transition of thermo-acoustic instability of supercritical hydrocarbon fuel flow". In: *Aerospace Science and Technology* 85 (2019), pp. 1–12. DOI: 10.1016/j.ast.2018.11.040.

[117] T. Kobayashi, S. Murayama, T. Hachijo, and H. Gotoda. "Early Detection of Thermoacoustic Combustion Instability Using a Methodology Combining Complex Networks and Machine Learning". In: *Physical Review Applied* 11.6 (2019), p. 064034. DOI: 10.1103/PhysRevApplied.11.064034.

[118] A.D. Fragkou, T.E. Karakasidis, and I.E. Sarris. "Recurrence quantification analysis of MHD turbulent channel flow". In: *Physica A: Statistical Mechanics and its Applications* (2019), p. 121741. DOI: 10.1016/j.physa.2019.121741.

[119] C. Liu, W.X. Zhou, and W.K. Yuan. "Statistical properties of visibility graph of energy dissipation rates in three-dimensional fully developed turbulence". In: *Physica A: Statistical Mechanics and its Applications* 389.13 (2010), pp. 2675–2681. DOI: 10.1016/j.physa.2010.02.043.

[120] Z. Gao, M. Small, and J. Kurths. "Complex network analysis of time series". In: *EPL (Europhysics Letters)* 116.5 (2017), p. 50001. DOI: 10.1209/0295-5075/116/50001.

[121] Y. Zou, R. V. Donner, N. Marwan, J. F. Donges, and J. Kurths. "Complex network approaches to nonlinear time series analysis". In: *Physics Reports* (2018). DOI: 10.1016/j.physrep.2018.10.005.

[122] B. Luque, L. Lacasa, F. Ballesteros, and J. Luque. "Horizontal visibility graphs: Exact results for random time series". In: *Physical Review E* 80.4 (2009), p. 046103. DOI: 10.1103/PhysRevE.80.046103.

[123] P. Manshour. "Complex network approach to fractional time series". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 25.10 (2015), p. 103105. DOI: 10.1063/1.4930839.

[124] A. M. Nuñez, L. Lacasa, J. P. Gomez, and B. Luque. "Visibility algorithms: A short review". In: *New Frontiers in Graph Theory* (2012), pp. 119–152. DOI: 10.5772/34810.

[125] X. Lan, H. Mo, S. Chen, Q. Liu, and Y. Deng. "Fast transformation from time series to visibility graphs". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 25.8 (2015), p. 083105. DOI: 10.1063/1.4927835.

[126] S. Ghosh and A. Dutta. "An efficient non-recursive algorithm for transforming time series to visibility graph". In: *Physica A: Statistical Mechanics and its Applications* 514 (2019), pp. 189–202. DOI: 10.1016/j.physa.2018.09.053.

[127] D. F. Yela, F. Thalmann, V. Nicosia, D. Stowell, and M. Sandler. "Efficient On-line Computation of Visibility Graphs". In: *arXiv preprint* (2019). URL: https://arxiv.org/abs/1905.03204.

[128] Iacobello Giovanni. *Fast natural visibility graph (NVG) for MATLAB*. MATLAB Central File Exchange. 2019. URL: www.mathworks.com/matlabcentral/fileexchange/70432-fast-natural-visibility-graph-nvg-for-matlab.

[129] Iacobello Giovanni. *Fast Horizontal Visibility Graph (HVG) for MATLAB*. MATLAB Central File Exchange. 2019. URL: www.mathworks.com/matlabcentral/fileexchange/72889-fast-horizontal-visibility-graph-hvg-for-matlab.

[130] S. Martin, W. M. Brown, R. Klavans, and K. W. Boyack. "OpenOrd: an open-source toolbox for large graph layout". In: *Visualization and Data Analysis*. Vol. 7868. (2011), p. 786806. DOI: 10.1117/12.871402.

[131] A. A. Tsonis and P. J. Roebber. "The architecture of the climate network". In: *Physica A: Statistical Mechanics and its Applications* 333 (2004), pp. 497–504. DOI: 10.1016/j.physa.2003.10.045.

[132] A. A. Tsonis, K. L. Swanson, and P. J. Roebber. "What do networks have to do with climate?" In: *Bulletin of the American Meteorological Society* 87.5 (2006), pp. 585–596. DOI: 10.1175/BAMS-87-5-585.

[133] K. Yamasaki, A. Gozolchiani, and S. Havlin. "Climate networks around the globe are significantly affected by El Nino". In: *Physical Review Letters* 100.22 (2008), p. 228501. DOI: 10.1103/PhysRevLett.100.228501.

[134] J. F. Donges, Y. Zou, N. Marwan, and J. Kurths. "Complex networks in climate dynamics". In: *The European Physical Journal Special Topics* 174.1 (2009), pp. 157–179. DOI: 10.1140/epjst/e2009-01098-2.

[135] S. Scarsoglio, F. Laio, and L. Ridolfi. "Climate dynamics: a network-based approach for the analysis of global precipitation". In: *PLoS One* 8.8 (2013), e71129. DOI: 10.1371/journal.pone.0071129.

[136] Y. Wang, A. Gozolchiani, Y. Ashkenazy, Y. Berezin, O. Guez, and S. Havlin. "Dominant imprint of Rossby waves in the climate network". In: *Physical Review Letters* 111.13 (2013), p. 138501. DOI: 10.1103/PhysRevLett.111.138501.

[137] J. F. Donges, I. Petrova, A. Loew, N. Marwan, and J. Kurths. "How complex climate networks complement eigen techniques for the statistical analysis of climatological data". In: *Climate dynamics* 45.9-10 (2015), pp. 2407–2424. DOI: 10.1007/s00382-015-2479-3.

[138] D. Zhou, A. Gozolchiani, Y. Ashkenazy, and S. Havlin. "Teleconnection paths via climate network direct link detection". In: *Physical Review Letters* 115.26 (2015), p. 268501. DOI: 10.1103/PhysRevLett.115.268501.

[139] Ankit Agarwal, Levke Caesar, Norbert Marwan, Rathinasamy Maheswaran, Bruno Merz, and Jürgen Kurths. "Network-based identification and characterization of teleconnections on different scales". In: *Scientific Reports* 9.1 (2019), p. 8808. DOI: 10.1038/s41598-019-45423-5.

[140] N. Boers, B. Goswami, A. Rheinwalt, B. Bookhagen, B. Hoskins, and J. Kurths. "Complex networks reveal global pattern of extreme-rainfall teleconnections". In: *Nature* 566.7744 (2019), p. 373. DOI: 10.1038/s41586-018-0872-x.

[141] E. Ser-Giacomi, V. Rossi, C. López, and Em. Hernández-García. "Flow networks: A characterization of geophysical fluid transport". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 25.3 (2015), p. 036404. DOI: 10.1063/1.4908231.

[142] L. Tupikina, N. Molkenthin, C. López, E. Hernández-García, N. Marwan, and J. Kurths. "Correlation networks from flows. The case of forced and time-dependent advection-diffusion dynamics". In: *PloS One* 11.4 (2016), e0153703. DOI: 10.1371/journal.pone.0153703.

[143] E. Ser-Giacomi, V. Rodriguez-Méndez, C. López, and E. Hernández-García. "Lagrangian Flow Network approach to an open flow model". In: *The European Physical Journal Special Topics* 226.9 (2017), pp. 2057–2068. DOI: 10.1140/epjst/e2017-70044-2.

[144] M. Gelbrecht, N. Boers, and J. Kurths. "A complex network representation of wind flows". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 27.3 (2017), p. 035808. DOI: 10.1063/1.4977699.

[145] M. Lindner and R. V. Donner. "Spatio-temporal organization of dynamics in a two-dimensional periodically driven vortex flow: A Lagrangian flow network perspective". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 27.3 (2017), p. 035806. DOI: 10.1063/1.4975126.

[146] N. Molkenthin, H. Kutza, L. Tupikina, N. Marwan, J. F. Donges, U. Feudel, J. Kurths, and R. V. Donner. "Edge anisotropy and the geometric perspective on flow networks". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 27.3 (2017), p. 035802. DOI: 10.1063/1.4971785.

[147] S. Fellini, P. Salizzoni, L. Soulhac, and L. Ridolfi. "Propagation of toxic substances in the urban atmosphere: A complex network perspective". In: *Atmospheric Environment* 198 (2019), pp. 291–301. DOI: 10.1016/j.atmosenv.2018.10.062.

[148] S. Scarsoglio, G. Iacobello, and L. Ridolfi. "Complex networks unveiling spatial patterns in turbulence". In: *International Journal of Bifurcation and Chaos* 26.13 (2016), p. 1650223. DOI: 10.1142/S0218127416502230.

[149] K. Taira, A. G. Nair, and S. L. Brunton. "Network structure of two-dimensional decaying isotropic turbulence". In: *Journal of Fluid Mechanics* 795 (2016). DOI: 10.1017/jfm.2016.235.

[150] P. S Krueger, M. Hahsler, E. V. Olinick, S. H. Williams, and M. Zharfa. "Quantitative classification of vortical flows based on topological features using graph matching". In: *Proceedings of the Royal Society A* 475.2228 (2019), p. 20180897. DOI: 10.1098/rspa.2018.0897.

[151] K. Caló, M. O. Khan, D. Gallo, S. Scarsoglio, D.A. Steinman, L. Ridolfi, and U. Morbiducci. "Exploring intracranial aneurysm hemodynamics with a complex networks approach". In: *Proceedings VII Meeting Italian Chapter of the European Society of Biomechanics* 1 (2017), pp. 28–29. URL: http://www.esb-ita.it/wp-content/uploads/2017/09/G15_Cal%C3%B2_ESBITA17.pdf.

[152] V. R. Unni, A. Krishnan, R. Manikandan, N. B. George, R.I. Sujith, N. Marwan, and J. Kurths. "On the emergence of critical regions at the onset of thermoacoustic instability in a turbulent combustor". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 28.6 (2018), p. 063125. DOI: 10.1063/1.5028159.

[153] S. Murayama, H. Kinugawa, I. T. Tokuda, and H. Gotoda. "Characterization and detection of thermoacoustic combustion oscillations based on statistical complexity and complex-network theory". In: *Physical Review E* 97.2 (2018), p. 022223. DOI: 10.1103/PhysRevE.97.022223.

[154] K. Takagi, H. Gotoda, T. Miyano, S. Murayama, and I. T. Tokuda. "Synchronization of two coupled turbulent fires". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 28.4 (2018), p. 045116. DOI: 10.1063/1.5009896.

[155] S. Murayama and H. Gotoda. "Attenuation behavior of thermoacoustic combustion instability analyzed by a complex-network-and synchronization-based approach". In: *Physical Review E* 99.5 (2019), p. 052222. DOI: 10.1103/PhysRevE.99.052222.

[156] T. Hashimoto, H. Shibuya, H. Gotoda, Y. Ohmichi, and S. Matsuyama. "Spatiotemporal dynamics and early detection of thermoacoustic combustion instability in a model rocket combustor". In: *Physical Review E* 99.3 (2019), p. 032208. DOI: 10.1103/PhysRevE.99.032208.

[157] A. Krishnan, R.I. Sujith, N. Marwan, and J. Kurths. "On the emergence of large clusters of acoustic power sources at the onset of thermoacoustic instability in a turbulent combustor". In: *Journal of Fluid Mechanics* 874 (2019), pp. 455–482. DOI: 10.1017/jfm.2019.429.

[158] H. A. Dijkstra, E. Hernández-García, C. Masoller, and M. Barreiro. "Computational Tools for Network Analysis". In: *Networks in Climate*. Cambridge University Press, (2019), pp. 79–93. DOI: 10.1017/9781316275757.006.

[159] A. Krishnan, P.R. Manikandan R .and Midhun, K.V. Reeja, V.R. Unni, R.I. Sujith, N. Marwan, and J. Kurths. "Mitigation of oscillatory instability in turbulent reactive flows: A novel approach using complex networks". In: *EPL (Europhysics Letters)* 128.1 (2019), p. 14003. DOI: 10.1209/0295-5075/128/14003.

[160] J. M. Wallace. "Space-time correlations in turbulent flow: A review". In: *Theoretical and Applied Mechanics Letters* 4.2 (2014), p. 022003. DOI: 10.1063/2.1402203.

[161] A. G. Nair and K. Taira. "Network-theoretic approach to sparsified discrete vortex dynamics". In: *Journal of Fluid Mechanics* 768 (2015), pp. 549–571. DOI: 10.1017/jfm.2015.97.

[162] M. G. Meena, A. G. Nair, and K. Taira. "Network community-based model reduction for vortical flows". In: *Physical Review E* 97.6 (2018), p. 063103. DOI: 10.1103/PhysRevE.97.063103.

[163] Z. Bai, N. B. Erichson, M. G. Meena, K. Taira, and S. L. Brunton. "Randomized methods to characterize large-scale vortical flow networks". In: *PLOS ONE* 14.11 (2019), pp. 1–19. DOI: 10.1371/journal.pone.0225265.

[164] A. Hadjighasem, D. Karrasch, H. Teramoto, and G. Haller. "Spectral-clustering approach to Lagrangian vortex detection". In: *Physical Review E* 93.6 (2016), p. 063107. DOI: 10.1103/PhysRevE.93.063107.

[165] K. L. Schlueter-Kuck and J. O. Dabiri. "Coherent structure colouring: identification of coherent structures from sparse data using graph theory". In: *Journal of Fluid Mechanics* 811 (2017), pp. 468–486. DOI: 10.1017/jfm.2016.755.

[166] K. L. Schlueter-Kuck and J. O. Dabiri. "Identification of individual coherent sets associated with flow trajectories using coherent structure coloring". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 27.9 (2017), p. 091101. DOI: 10.1063/1.4993862.

[167] C. Schneide, A. Pandey, K. Padberg-Gehle, and J. Schumacher. "Probing turbulent superstructures in Rayleigh-Bénard convection by Lagrangian trajectory clusters". In: *Physical Review Fluids* 3.11 (2018), p. 113501. DOI: 10.1103/PhysRevFluids.3.113501.

[168] R. Chakraborty, A. Coutino, and M. Stastna. "Particle clustering and subclustering as a proxy for mixing in geophysical flows". In: *Nonlinear Processes in Geophysics* 26.3 (2019), pp. 307–324. DOI: 10.5194/npg-26-307-2019.

[169] K. L. Schlueter-Kuck and J. O. Dabiri. "Model parameter estimation using coherent structure colouring". In: *Journal of Fluid Mechanics* 861 (2019), pp. 886–900. DOI: 10.1017/jfm.2018.898.

[170] K. Padberg-Gehle and C. Schneide. "Network-based study of Lagrangian transport and mixing". In: *Nonlinear Processes in Geophysics* 24.4 (2017), p. 661. DOI: 10.5194/npg-24-661-2017.

[171] I. I. Rypina and L. J. Pratt. "Trajectory encounter volume as a diagnostic of mixing potential in fluid flows". In: *Nonlinear Processes in Geophysics* 24.2 (2017), pp. 189–202. DOI: 10.5194/npg-24-189-2017.

[172] B. E. Husic, K. L. Schlueter-Kuck, and J. O. Dabiri. "Simultaneous coherent structure coloring facilitates interpretable clustering of scientific data by amplifying dissimilarity". In: *PloS One* 14.3 (2019), e0212442. DOI: 10.1371/journal.pone.0212442.

[173] G. Iacobello, L. Ridolfi, M. Marro, P. Salizzoni, and S. Scarsoglio. "Complex Network Analysis of Wind Tunnel Experiments on the Passive Scalar Dispersion in a Turbulent Boundary Layer". In: *Progress in Turbulence VIII*. Ed. by R. Örlü, A. Talamelli, J. Peinke, and M. Oberlack. Cham: Springer International Publishing, (2019), pp. 215–220. DOI: 10.1007/978-3-030-22196-6_34.

[174] B. Luque, F. J. Ballesteros, A. Robledo, and L. Lacasa. "Entropy and renormalization in chaotic visibility graphs". In: *Mathematical Foundations and Applications of Graph Entropy*. Vol. 6. Wiley Online Library, (2016), pp. 1–39. DOI: 10.1002/9783527693245.ch1.

[175] E. Zhuang, M. Small, and G. Feng. "Time series analysis of the developed financial markets' integration using visibility graphs". In: *Physica A: Statistical Mechanics and its Applications* 410 (2014), pp. 483–495. DOI: 10.1016/j.physa.2014.05.058.

[176] U. Hasson, J. Iacovacci, B. Davis, R. Flanagan, E. Tagliazucchi, H. Laufs, and L. Lacasa. "A combinatorial framework to quantify peak/pit asymmetries in complex dynamics". In: *Scientific reports* 8.1 (2018), p. 3557. DOI: 10.1038/s41598-018-21785-0.

[177] R. V. Donner and J. F. Donges. "Visibility graph analysis of geophysical time series: Potentials and possible pitfalls". In: *Acta Geophysica* 60.3 (2012), pp. 589–623. DOI: 10.2478/s11600-012-0032-x.

[178] A. W. Vreman and J. G. M. Kuerten. "Comparison of direct numerical simulation databases of turbulent channel flow at $\mathrm{Re}_\tau = 180$". In: *Physics of Fluids* 26 (2014), p. 015102. DOI: 10.1063/1.4861064.

[179] Maurizio Quadrio and Paolo Luchini. "Integral space–time scales in turbulent wall flows". In: *Physics of fluids* 15.8 (2003), pp. 2219–2227. DOI: 10.1063/1.1586273.

[180] C Xu, Z Zhang, J MiJ den Toonder, and FTM Nieuwstadt. "Origin of high kurtosis levels in the viscous sublayer. Direct numerical simulation and experiment". In: *Physics of Fluids* 8.7 (1996), pp. 1938–1944. DOI: 10.1063/1.868973.

[181] S. P. Arya. *Air pollution meteorology and dispersion*. Oxford University Press New York, (1999), p. 310.

[182] M. Tumminello, T. Aste, T. Di Matteo, and R. N. Mantegna. "A tool for filtering information in complex systems". In: *Proceedings of the National Academy of Sciences* 102.30 (2005), pp. 10421–10426. DOI: 10.1073/pnas.0500298102.

[183] G. I. Taylor. "The spectrum of turbulence". In: *Proceedings of the Royal Society of London A* 164.919 (1938), pp. 476–490. DOI: 10.1098/rspa.1938.0032.

[184] C. Geng, G. He, Y. Wang, C. Xu, A. Lozano-Durán, and J. M. Wallace. "Taylor's hypothesis in turbulent channel flow considered using a transport equation analysis". In: *Physics of Fluids* 27.2 (2015), p. 025111. DOI: 10.1063/1.4908070.

[185] R. N. Mantegna. "Hierarchical structure in financial markets". In: *The European Physical Journal B* 11.1 (1999), pp. 193–197. DOI: 10.1007/s100510050929.

[186] Mikkel Thorup. "Undirected single-source shortest paths with positive integer weights in linear time". In: *Journal of the ACM (JACM)* 46.3 (1999), pp. 362–394. DOI: 10.1145%2F316542.316548.

[187] G. Iacobello, S. Scarsoglio, J.G.M. Kuerten, and L. Ridolfi. "Lagrangian network analysis of turbulent mixing". In: *Journal of Fluid Mechanics* 865 (2019), pp. 546–562. DOI: 10.1017/jfm.2019.79.

[188] H. B. Fischer. "Longitudinal dispersion and turbulent mixing in open-channel flow". In: *Annual Review of Fluid Mechanics* 5.1 (1973), pp. 59–78. DOI: 10.1146/annurev.fl.05.010173.000423.

[189] H. J. S. Fernando. *Handbook of Environmental Fluid Dynamics, Volume Two: Systems, Pollution, Modeling, and Measurements*. CRC press Taylor & Francis Group, (2012). DOI: 10.1201/b13691.

[190] J. W. Elder. "The dispersion of marked fluid in turbulent shear flow". In: *Journal of Fluid Mechanics* 5.4 (1959), pp. 544–560. DOI: 10.1017/S0022112059000374.

[191] B. J. Geurts and J. G. M. Kuerten. "Ideal stochastic forcing for the motion of particles in large-eddy simulation extracted from direct numerical simulation of turbulent channel flow". In: *Physics of Fluids* 24 (2012), p. 081702. DOI: 10.1063/1.4745857.

[192] T. Opsahl, F. Agneessens, and J. Skvoretz. "Node centrality in weighted networks: Generalizing degree and shortest paths". In: *Social Networks* 32.3 (2010), pp. 245–251. DOI: 10.1016/j.socnet.2010.03.006.

[193] Y. Li, E. Perlman, M. Wan, Y. Yang, C. Meneveau, R. Burns, S. Chen, A. Szalay, and G. Eyink. "A public turbulence database cluster and applications to study Lagrangian evolution of velocity increments in turbulence". In: *Journal of Turbulence* 9 (2008), N31. DOI: 10.1080/14685240802376389.

[194] J. Graham, K. Kanov, X.I.A. Yang, M. Lee, N. Malaya, C.C. Lalescu, R. Burns, G. Eyink, A. Szalay, R.D. Moser, and C. Meneveau. "A web services accessible database of turbulent channel flow and its use for testing a new integral wall model for LES". In: *Journal of Turbulence* 17.2 (2016), pp. 181–215. DOI: 10.1080/14685248.2015.1088656.

[195] P. R. Spalart, R. D. Moser, and M. M. Rogers. "Spectral methods for the Navier-Stokes equations with one infinite and two periodic directions". In: *Journal of Computational Physics* 96 (1991), pp. 297–324. DOI: 10.1016/0021-9991(91)90238-G.

[196] W. R. Michałek, J. G. M. Kuerten, J. C. H. Zeegers, R. Liew, J. Pozorski, and B. J. Geurts. "A hybrid stochastic-deconvolution model for large-eddy simulation of particle-laden flow". In: *Physics of Fluids* 25 (2013), p. 123302. DOI: 10.1063/1.4849536.

[197] A. W. Vreman and J. G. M. Kuerten. "Statistics of spatial derivatives of velocity and pressure in turbulent channel flow". In: *Physics of Fluids* 26 (2014), p. 085103. DOI: 10.1063/1.4891624.

[198] H.P.A.H. Irwin. "The design of spires for wind simulation". In: *Journal of Wind Engineering and Industrial Aerodynamics* 7.3 (1981), pp. 361–366. DOI: 10.1016/0167-6105(81)90058-1.

[199] J.E. Fackrell. "A flame ionisation detector for measuring fluctuating concentration". In: *Journal of Physics E: Scientific Instruments* 13.8 (1980), p. 888. DOI: 10.1088/0022-3735/13/8/021.

[200] Finn E Jørgensen. *How to measure turbulence with hot-wire anemometers - a practical guide, Dantec dynamics*. 2002. URL: www.dantecdynamics.com/how-to-measure-turbulence-practical-guide.

This Ph.D. thesis has been typeset by means of the TEX-system facilities. The typesetting engine was pdfLATEX. The document class was `toptesi`, by Claudio Beccari, with option `tipotesi=scudo`. This class is available in every up-to-date and complete TEX-system installation.